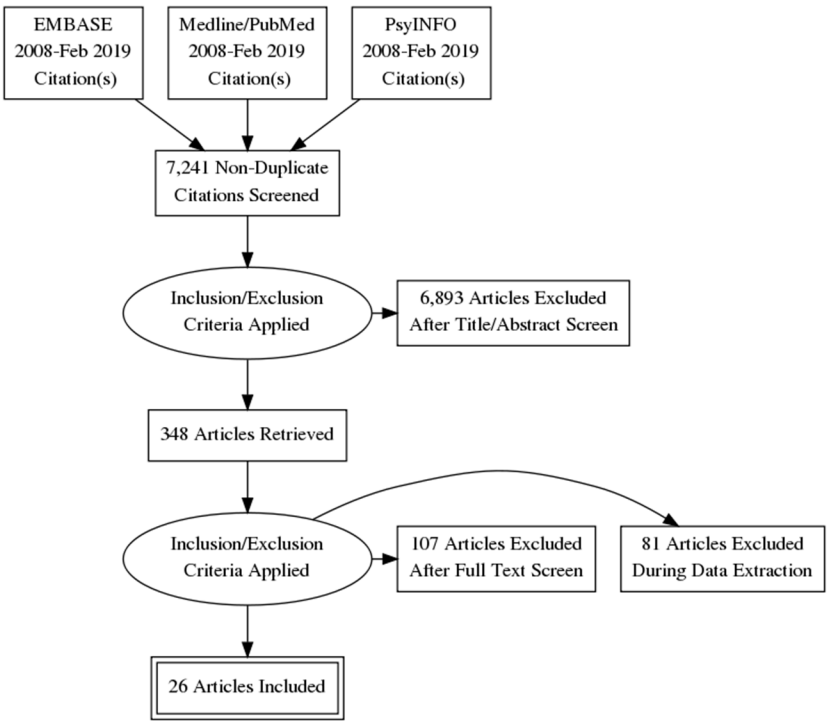
Chapter 4: Complementary and Integrative Health (CIH) and other interventions for Treating Major Depressive Disorder (MDD)

Results of the Literature Search for MDD

Extensive literature searches identified 7,241 citations (after duplicates removed) potentially addressing the CIH and other interventions of interest for the treatment of MDD or for individuals at risk of suicide. Of those, 6,893 were excluded upon title and abstract review for clearly not meeting inclusion criteria (e.g., not pertinent to the topic, not published in English, published prior to study inclusion publication date, or not a full-length article). A total of 348 full-length articles were retrieved for review (See Error! Reference source not found. for the PRISMA diagram). Of those, 107 were excluded due to having the wrong patient population (27 studies), the wrong study design (26 studies), the wrong intervention (24 studies), wrong outcomes (13 studies), duplicates (4 studies), all studies included in systematic review were published prior to 2008 (3 studies), conference abstracts (3 studies), less than 20 patients (2 studies), more recent and/or comprehensive systematic review available (2 studies), wrong comparator (2 studies that compared different intensity level of exercise), and protocol (1 study). An additional 81 studies were excluded during data abstraction. Reasons for these exclusions are listed in **Appendix B**.

Figure 1. Prisma Study Flow Diagram for MDD



Overall, 26 studies were included in the systematic review for MDD. **Table 1** presents a summary of the evidence (how many RCTs and/or SRs) for each CIH intervention.

Table 1. Overview of Evidence for CIH and other Non-Concentional Interventions to Treat Major Depressive Disorder

Intervention	Number of Studies and Types of Studies	Strength of Evidence (SOE)
Accelerated Resolution Therapy (ART)	0	NA
Acupuncture	1 SR (with 64 RCTs)	Very low to Moderate
Art therapy	1 SR (with 2 RCTs); 2 RCTs	Very low to Low
Cannabinoids	0	NA
Chiropractic care	0	NA
Equine therapy	0	NA
Exercise therapy (outdoor therapy)[1]	1 SR (with 23 RCTs); 9 RCTs	Very low to Low
Healing Touch	0	NA
Hyperbaric Oxygen Therapy	0	NA
Massage therapy	0	NA
Meditation	3 RCTs	Very low to Low
Music therapy	1 RCT	Very low to Low
Relaxation therapy	0	NA
Tai chi	3 RCTs	Low
Therapeutic touch	0	NA
Training and caring for service dogs	0	NA
Transcranial Magnetic Stimulation (TMS)	2 SRs (with 104 RCTs); 2 RCTs	Low to Moderate
Yoga	1 SR (with 11 RCTs)	Very low
Total Studies	**26 studies (6 SRs with 204 RCTs and 20 additional RCTs)**	

RCT: Randomized controlled trial; SR: systematic review

All of the full-text studies included in this report along with further details of the search terms and concepts used to guide the searches for MDD are provided in a supplemental file on Max.gov and can be accessed here: https://community.max.gov/display/VAExternal/MDD+Report+Supplementary+Materials

[1] It is important to note that types of exercise vary across studies and conditions. Outdoor therapy was identified in the CARA legislation, while exercise was identified by the COVER Commission as an intervention of interest. These have been combined due to the overlap in the studies.

Acupuncture

Evidence Base

Our searches of the literature identified 1 SR with an evidence base of 64 RCTs published between 1985 to 2016 that addressed the use of acupuncture to treat adults diagnosed with depression. Smith (2018) examined acupuncture compared to medication (42 RCTs), control acupuncture, which is a treatment that looks similar to active acupuncture (14 RCTs), no treatment (5 RCTs), or psychological therapies including counseling and an educational intervention with psychological guidance (2 RCTs). Control acupuncture may include an invasive acupuncture control, sham acupuncture, which involves the insertion of a needle into a non-acupuncture site, minimal acupuncture in which needles are inserted into non-acupuncture sites in a superficial way to avoid stimulation or manipulation, non-invasive actupuncture control with a placbo needle, mock elcetro-acupuncutre with decommissioned acupuncture stimulation, or mock laser acupuncture. The studies in this review included a total of 7,104 adults with clinical depression and considered the efficacy of manual acupuncture (42 RCTs), electroacupuncture (13 RCTs), a combination of manual and electroacupuncture (7 RCTs), and laser acupuncture (2 RCTs). Treatment sessions ranged from <10 sessions to 60 sessions, with an average of 30 total sessions, lasting from 20 to 60 minutes of needling per session. The primary outcome measured was reduction in the severity of depression at the end of treatment. Secondary outcomes included remission, quality of life, and adverse events.

Study Quality

Using the AMSTAR instrument, we rated the quality of the systematic review as high (See **Table 4** for more information on the review ratings). The authors of the review used the Cochrane tool to assess the RoB of the included studies. The trials were rated as low to high RoB with studies of high RoB downgraded due to lack of blinding patients, study staff, and/or outcome assessors.

Key Findings

Below, we describe the key findings for the outcomes of interest with the GRADE strength of the evidence (SOE) rating. See **Table 1** for factors that influenced the SOE ratings.

Acupuncture vs. No Treatment/Wait List/Treatment as Usual

- Combined evidence of 5 RCTs suggests that acupuncture (manual and electro) may moderately reduce depression severity at posttreatment compared to no treatment/wait list/treatment as usual (SOE: Low).
- Combined evidence from 2 RCTs found no difference between acupuncture and manual acupuncture in remission of depression (SOE: Moderate).
- Evidence from 1 RCT found no difference between acupuncture and no treatment in the risk of adverse events (SOE: Moderate).

Acupuncture vs. Control Acupuncture (invasive or non-invasive sham controls)

- Combined evidence of 14 RCTs suggests that acupuncture may be associated with a small reduction in depression severity of 1.69 points on the Hamilton Depression Rating Scale (HAMD) at post-treatment (SOE: Low).

- Combined evidence of 10 RCTs found greater remission following acupuncture when compared to control acupuncture (SOE: Moderate).
- Combined evidence of 5 RCTs found no clear between group differences in the risk of adverse events (SOE: Moderate).
- Combined evidence of 2 RCTs found no clear between group differences in emotional quality of life (SOE: Moderate).
- Evidence from 1 RCT found no clear between group differences in physical quality of life (SOE: Moderate).

Acupuncture vs. Medication

- Combined evidence of 31 RCTs suggest that acupuncture may lead to a small reduction in depression severity compared to medication alone at post-treatment (SOE: Very low).
- Combined evidence from 27 RCTs suggest remission from acupuncture compared to medication alone (SOE: Moderate).
- Combined evidence of 3 RCTs suggest lower ratings of adverse events following acupuncture compared with medication alone (SOE: Very low).

Acupuncture + Medication vs. Medication Alone

- Combined evidence from 11 RCTs suggest that acupuncture plus medication statistically significantly reduces depression severity compared to medication alone post-treatment. (SOE: Very low).
- Combined evidence from 9 RCTs show no clear difference in remission between acupuncture used in conjunction with medication compared to medication alone (SOE: Low).
- Combined evidence from 3 RCTs show no clear difference in adverse events associated with manual and/or electro-acupuncture plus medication (SSRIs) compared to medication (SSRIs) alone (SOE: Very low).
- Combined evidence from 2 RCTs found no clear between group differences in emotional quality of life (SOE: Low).
- Evidence from 1 RCT found suggests that acupuncture plus medication may improve physical quality of life compared to medication alone (SOE: Low).

Acupuncture vs. Psychological Treatment

- Combined evidence of 2 RCTs show no clear differences between acupuncture and psychological therapy in depression severity at post-treatment (SOE: Low).
- Evidence from 1 RCT suggests no differences between groups in rates of adverse events (SOE: Low).

Discussion

The evidence for acupuncture in the treatment of clinical depression shows there may be a moderate reduction in the severity of depression when compared with treatment as usual or no treatment. The use of acupuncture may lead to a small reduction in the severity of depression when compared with control acupuncture. Acupuncture given alone or as an adjunct to medication may reduce depression symptoms severity, however, it is important to note that the quality of the evidence is very low. No significant difference was observed between acupuncture and psychological treatment in the reduction of depression symptoms. Studies show substantial variation resulting from use of different classes of medications and different modes of acupuncture stimulation. Most included studies did not report adverse events; therefore, it is unclear what risks of adverse events there are with acupuncture. Lack of medium and long-term follow-up in clinical trials represents a significant limitation of the evidence base.

Table 1. Strength of Evidence for Acupuncture to Treat MDD

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
Acupuncture vs. No Treatment/Wait List/Treatment as Usual									
Depression	5 RCTs in Smith (2018)	ACU vs. No tx/WL/TAU (n=488)	SMD: -0.66, 95% CI -1.06 to -0.25, p=0.0014; favors ACU	Yes (-1)	Yes (-1); substantial heterogeneity	No	No	No	Low
Remission	2 RCTs in Smith (2018)	ACU vs Manual ACU (n=94)	RR: 1.67; 95% CI 0.77 to 3.65, p=0.20; NS	Yes (-1)	No	No	No	No	Moderate
Adverse events	1 RCT in Smith (2018)	ACU vs. No tx/WL/TAU (n=302)	RR: 0.89; 95% CI 0.35 to 2.24, p=0.80; NS	No	No	No	Yes (-1); small sample size	No	Moderate
Acupuncture vs. Control Acupuncture (invasive, non-invasive sham controls)									
Depression	14 RCTs in Smith (2018)	ACU vs. Control ACU (n=841)	SMD: -1.69; 95% CI -3.33 to -0.05, p=0.043; favors ACU	Yes (-1)	Yes (-1); substantial heterogeneity	No	No	No	Low
Remission	10 RCTs in Smith (2018)	ACU vs Control ACU (n=601)	RR: 1.91; 95% CI 1.14 to 3.21;	Yes (-1)	No	No	No	No	Moderate

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
			p=0.00024, favors ACU						
Adverse events	5 RCTs in Smith (2018)	ACU vs. Control ACU (n=300)	RR: 1.63; 95% CI 0.93 to 2.86, p=0.087; NS	No	No	No	Yes (-1); small sample size	No	Moderate
Quality of life (emotional)	2 RCTs in Smith (2018)	ACU vs. Control ACU (n=167)	SMD: -2.25; 95% CI -5.89 to 1.39, p=0.23; NS	No	No	No	Yes (-1); small sample size	No	Moderate
Quality of life (physical)	1 RCT in Smith (2018)	ACU vs. Control ACU (n=150)	SMD: -5.12; 95% CI -10.38 to 0.13, p=0.056; NS	No	No	No	Yes (-1); 1 small study	No	Moderate
Acupuncture vs. Medication									
Depression	31 RCTs in Smith (2018)	ACU vs. Med (n=3,127)	SMD: -0.23; 95% CI -0.40 to -0.05, p=0.011; favors ACU	Yes (-1)	Yes (-2); considerable heterogeneity	No	No	No	Very low
Remission	27 RCTs in Smith (2018)	ACU vs. Med (n=2,918)	RR: 1.16; 95% CI 1.05 to	Yes (-1)	No	No	No	No	Moderate

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
			1.29; p=0.004, favors ACU						
Adverse events	3 RCTs in Smith (2018)	ACU vs. Med (n=481)	SMD: -4.32; 95% CI -7.41 to -1.23, p=0.0061; favors ACU	Yes (-1)	Yes (-2); considerable heterogeneity	No	No	No	Very low
Acupuncture + Medication vs. Medication Alone									
Depression	11 RCTs in Smith (2018)	ACU + Med vs. Med (n=813)	SMD: -1.15; 95% CI -1.63 to -0.66, p<.00001; favors ACU+Med	Yes (-1)	Yes (-2); considerable heterogeneity	No	No	No	Very low
Remission	9 RCTs in Smith (2018)	ACU + Med vs. Med (n=618)	RR: 1.21; 95% CI 0.85 to 1.73, p=0.29, NS	Yes (-1)	Yes (-1); substantial heterogeneity	No	No	No	Low
Quality of life (physical)	1 RCT in Smith (2018)	ACU + Med vs. Med (n=127)	SMD: 1.19; 95% CI 0.33 to 2.05, p=0.0066; favors ACU + Med	Yes (-1)	No	No	Yes (-1); small sample size	No	Low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
Quality of life (emotional)	2 RCTs in Smith (2018)	ACU + Med vs. Med (n=219)	SMD: 0.25; 95% CI -0.90 to 1.40, p=0.67; NS	Yes (-1)	Yes (-1); substantial heterogeneity	No	No	No	Low
Adverse events	3 RCTs in Smith (2018)	ACU + Med vs. Med (n=200)	SMD: -1.32; 95% CI -2.86 to 0.23, p<0.095; NS	Yes (-1)	Yes (-2); considerable heterogeneity	No	No	No	Very low
Acupuncture vs. Psychological Therapy									
Depression	2 RCTs in Smith (2018)	ACU vs. Psych Therapy (n=497)	SMD: -0.50; 95% CI -1.33 to 0.33, p=0.24; NS	Yes (-1)	Yes (-1); substantial heterogeneity	No	No	No	Low
Adverse events	1 RCT in Smith (2018)	ACU vs. Psych Therapy (n=453)	RR: 0.62; 95% CI 0.29 to 1.33, p=0.22; NS	Yes (-1)	No	No	Yes (-1); only 1 study	No	Low

ACU: acupuncture; CI: confidence interval; CT: control group; ES: effective size; Med.: medication; mos.: months; NR: not reported; NS: not significant; RCT: randomized controlled trials; SE: standard error; SMD: standardized mean difference; TAU: treatment as usual; WL: waitlist

Table 2. GRADE Factors Used to Assess the Quality of a Body of Evidence

Evidence Category	Definition
Study Quality (Internal Validity or Risk of Bias)	Study quality considers the overall risk of bias rating of all the studies included in the evidence base. In this review, the overall risk of bias would be the average or median USPSTF rating for studies comprising an evidence base for a key outcome.
Consistency of Evidence	Consistency of evidence refers to the degree of similarity in the direction of effects or the degree of similarity in the effect sizes (magnitude of effect) across individual studies within an evidence base.
Directness of Evidence	Direct evidence directly compares interventions of interest in populations of interest and measures patient-oriented outcomes. Evidence can be indirect if the tested intervention differs from the intervention of interest, the study population differs from the population of interest, the outcomes differ from those of primary interest, or treatment comparisons have not been tested in head-to-head comparisons.
Precision of Evidence	Precision is the degree of certainty surrounding an estimate of effect with respect to an outcome. Precision is primarily assessed by examining the 95% confidence intervals around the summary effect size.

Link to GRADE Handbook: **http://gdt.guidelinedevelopment.org/app/handbook**

Table 3. Evidence Table for Systematic Review on Acupuncture to Treat MDD

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
Reference: Smith et al. 2018 **Organization/Country:** The Cochrane Collaboration, New South Wales, Australia **Purpose:** To examine the effectiveness and adverse effects of acupuncture for treatment of individuals with depression. **AMSTAR Rating: High** **Overall RoB of Included Studies:** Low to High using Cochrane tool. Methodological limitation include performance bias and incomplete data.	**Databases Searched:** Cochrane Common Mental Disorders Group Controlled Trials Register (CCMD-TR); Korean Studies Information Service System (KISS); DBPIA; Korea Institute of Science and Technology Information; Research Information Service System (RISS); Korea Med; Korean Medical Databases (KM base); Oriental Medicine Advanced Searching Integrated System (OASIS), several Korean medical journals. **Dates Searched:** Inception to June 2016 **Inclusion/Exclusion Criteria:** Published and unpublished RCTs with adult pts. aged ≥16 yrs. with a clinical diagnosis of depression using valid diagnostic measures. Studies must have assessed acupuncture compared to control acupuncture, no treatment, medication, other structured psychotherapies, or standard care. Included studies of acupuncture, electro-acupuncture, and laser acupuncture. **Final Evidence Base:** 64 RCTs	**Diagnosis:** Clinical depression **Number of Patients:** 7,104 **Age:** ≥16 yrs. or older **Gender:** Male and female	**Intervention:** Manual acupuncture (MA, 42 RCTs), electroacupuncture (13 RCTs), manual + electroacupuncture (7 RCTs), and laser acupuncture (2 RCTs). Treatment sessions varied from <10 to 60 sessions lasting from 20 to 60 min. **Comparators:** Medication (42 RCTs), control acupuncture (14 RCTs), no treatment (5 RCTs), psychological therapy (2 RCTs) **Follow-up:** During treatment, post-intervention, 0-6 months (short-term), 6-12 months (medium-term), >12 months (long-term) **Outcomes:** Primary outcomes: reduction in depression severity, adverse events. Secondary outcomes: remission of depression, QoL, change in use of medication or other support systems, dropouts from treatment	**ACU vs. MED** **Depression:** 31 RCTs (n=3,127), SMD: -0.23, 95% CI -0.40 to -0.05, I^2=80%; favors ACU **Remission:** 27 RCTs (n=2,918), RR: 1.16; 95% CI 1.05 to 1.29, I^2=24%; favors ACU **Adverse events:** 3 RCTs (n=481), SMD: -4.32; 95% CI -7.41 to -1.23, I^2=97%; favors ACU **ACU + MED vs. MED** **Depression:** 11 RCTs (n=813), SMD: -1.15, 95% CI -1.63 to -0.66, I^2=89%; favors ACU **Remission:** 9 RCTs (n=127), RR: 1.21; 95% CI 0.85 to 1.73, I^2=61%; NS **Adverse events:** 3 RCTs (n=200), SMD: -1.32; 95% CI -2.86 to 0.23, I^2=95%; favors ACU **QoL (physical):** 1 RCT (n=127), SMD: 1.19; 95% CI 0.33 to 2.05, I^2=0%; favors ACU+Med **QoL (emotional):** 2 RCTs (n=219), SMD: 0.25; 95% CI -0.90 to 1.40, I^2=71%; favors control **ACU vs. Control ACU** **Depression:** 14 RCTs (n=841), SMD: -1.69, 95% CI -3.33 to -0.05, I^2=80%; favors ACU

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
				Remission: 10 RCTs (n=601), RR: 1.91, 95% CI 1.14 to 3.21, I^2=48%, favors ACU **Adverse events:** 5 RCTs (n=300), RR: 1.63, 95% CI 0.93-2.86, I^2=10%; favors control ACU **QoL (emotional):** 2 RCTs (n=167), SMD: -2.25; 95% CI -5.89 to 1.39, I^2=0%; favors control ACU **QoL (physical):** 1 RCT (n=150), SMD: -5.12; 95% CI -10.38 to 0.13, I^2=52%; favors control ACU **ACU vs. No Treatment** **Depression:** 5 RCTs (n=488), SMD: -0.66, 95% CI -1.06 to -0.25, I^2=64%; favors ACU **Remission:** 2 RCTs (n=94), RR: 1.67, 95% CI 0.77 to 3.65, I^2=0%, NS **Adverse events:**1 RCT (n=302), RR: 0.89, 95% CI 0.35-2.24, favors no treatment **ACU vs. Psychotherapy** **Depression:** 2 RCTs (n=497), SMD: -0.50, 95% CI -1.33 to 0.33, I^2=85%; favors ACU **Adverse events**: 1 RCT (n=453), RRL 0.62, 95% CI 0.29-1.33, favors ACU Meta-regression results suggest moderate reduction in depression severity in favor of ACU

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
				compared with TAU/no tx and small reduction in depression severity in favor of ACU compared with control ACU. Effects of ACU vs. Med and Psychological therapy are uncertain due to very low quality of evidence. Risks of AEs w/ ACU are unclear as most studies have not reported on them. **Limitations:** Lack of medium and long-term f/u in clinical trials

ACU: acupuncture; AEs: adverse events; CI: confidence interval; CT: control group; ES: effective size; F/u: follow-up; I^2: % of heterogeneity between studies; Med: medication; mos.: months; NR: not reported; NS: not significant; QoL: quality of life; RCT: randomized controlled trials; SE: standard error; SMD: standardized mean difference; TAU: treatment as usual; Tx: treatment; WL: waitlist

Table 4. Systematic Review Risk of Bias AMSTAR Checklist Table on Acupuncture to Treat MDD

Question	Smith et al. (2018)
Did the research questions and inclusion criteria for the review include the components of PICO?	Yes
Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	Yes
Did the review authors explain their selection of the study designs for inclusion in the review?	Yes
Did the review authors use a comprehensive literature search strategy?	Yes
Did the review authors perform study selection in duplicate?	Yes
Did the review authors perform data extraction in duplicate?	Yes
Did the review authors provide a list of excluded studies and justify the exclusions?	Yes
Did the review authors describe the included studies in adequate detail?	Yes
Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	Yes
Did the review authors report on the sources of funding for the studies included in the review?	No
If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results? RCTs?	Yes
If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	Yes
Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	Yes
Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	Yes
If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	Yes
Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	Yes
Overall Quality	**High**

RoB: risk of bias

Table 5. AMSTAR Rating of Overall Confidence in Results of the Review

Category	Definition
High	*No or one non-critical weakness*: the systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest.
Moderate	*More than one non-critical weakness*: the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review.
Low or Very Low	*One or more critical flaw(s) with or without non-critical weaknesses*: the systematic review has one or more critical flaws and may not provide an accurate and comprehensive summary of the available studies that address the question of interest.

AMSTAR checklist, go to https://amstar.ca/Amstar_Checklist.php

References

Smith, C., Armour, M., Lee, M., Wang, L., & Hay, P. (2018). Acupuncture for depression. *Cochrane Database of Systematic Reviews, 3*(CD004046), 1-225.

Art Therapy

Evidence Base

Our searches of the literature identified 1 SR and 2 RCTs that assessed the use of art therapy in the treatment of adults diagnosed with depression. See **Table 3** and **Table 6** for details about the patients, interventions, outcomes and findings of the identified studies.

In brief, Meekums et al. (2015) conducted a SR to examine the effectiveness of Dance Movement Therapy (DMT) for depression with or without standard care, compared to no treatment or standard care alone, psychological therapies, pharmacological therapies, or other physical interventions including exercise and dance (Meekums et al., 2015). The authors also compared the effectiveness of different DMT approaches. The evidence base for the SR included a total of 3 RCTs enrolling 147 participants. However, only 2 of the RCTs (n=107) enrolled adults and will therefore, be included in this review.

One additional RCT by Blomdahl et al. (2018) randomized 79 patients with moderate to severe depression to receive either a manual-based phenomenological art therapy plus treatment as usual (PATd/TAU) (n=43) or TAU alone (n=36). According to Blomdahl et al. (2018), phenomenology art therapy focuses how someone perceives the world, their lives, and themselves with the aim of increasing self-awareness and understanding, accepting one's strengths and limits, and learn to prioritize these based on self-knowledge. Treatment as usual is this study mostly included various forms of pharmacotherapy and psychotherapy, but also include acupuncture for three of the participants. Treatment as usual was determined and performed by participants' regular physicians or therapists. Participation in the study did not affect the content or number of treatments in TAU in either group, rather the authors added TAU to PATd in order to test its impact. PATd consisted of 10, 60-minute long treatment sessions delivered weekly and both groups received the same number of TAU. The primary outcome of interest includes depression levels, and the secondary outcome of interest includes suicide ideation.

The other RCT, Ciasca et al. (2018) randomized 56 elderly women with MDD and were stable on pharmacotherapy to art therapy (n=31) or the control group which did not receive any adjunctive intervention. The art therapy group received 20, 90-minute weekly sessions provided by an art therapist. The primary outcome of interest includes depression levels.

Study Quality

Using the AMSTAR instrument, we rated the quality of the Meekums et al. (2015) review as high (see **Table 4** for the quality ratings). The authors of the review by Meekums rated the RoB of the included RCTs as moderate to high using criteria from the Cochrane tool. The authors indicated that some of the studies did not blind patients, clinicians, or outcome assessors. We rated the RoB of the Blomdahl and Ciasca trial as having some concerns due to lack of information on blinding of the outcome assessors and lack of information on the randomization process used (see **Table 7** for individual quality ratings).

Key Findings

Below, we describe the key findings for the outcomes of interest with the GRADE strength of the evidence (SOE) rating. See **Table 1** for factors that influenced the SOE ratings.

- Evidence from 2 RCTs suggest that Dance Movement Therapy led to a reduction in depression compared to standard care, however the findings did not reach clinical significance (SOE: Low).
- Evidence from 1 RCT suggests that there is no statistically significant difference between Dance Movement Therapy and standard care in improving quality of life (SOE: Very low).
- Evidence from 1 RCT suggests that PATd in conjunction with TAU resulted in a statistically significantly reduction in depression symptoms compared with TAU alone (SOE: Very low)
- Evidence from 1 RCT suggests that there is no statistically significant difference between PATd in conjunction with TAU compared to TAU alone in change in suicide intentions (SOE: Very low).
- Evidence from 1 RCT suggests that art therapy as an adjunct to pharmacotherapy in older adults statistically significantly improved depressive symptoms when compared to pharmacotherapy alone (SOE: Low).
- Evidence from 1 RCT suggests that art therapy in conjunction with pharmacotherapy in older adults statistically significantly improved symptoms of anxiety when compared to pharmacotherapy alone (SOE: Low).

Discussion

Overall, the results of the Meekums review suggests that, while there was evidence of a reduction in depression for group Dance Movement Therapy (DMT) conducted over a period between 4 and 10 weeks with a total of 20 sessions and combined with standard care vs. standard care alone, the result was not statistically significant (standardized mean difference [SMD]: -7.33; 95% [confidence interval] CI -9.92 to -4.73). For quality of life, one study showed no effect in either direction (SMD: 0.30, 95% CI -0.60 to 1.20).

Blomdahl et al. RCT suggests that manual-based Art Therapy in addition to treatment as usual (PATd/TAU) demonstrates a statistically significant difference for improvement in depression levels (MADRS-S) compared to treatment as usual (TAU) alone. Those in the art therapy group (n=43) received 10, 1-hour weekly sessions including various tasks that served as prompts for the participant to paint. This study also examined suicide ideation as a secondary outcome, but no statistically significant difference was found in suicide ideation between the PATd/TAU group and the TAU group.

Ciasca et al. RCT showed statistically significant improvements in depression levels (GDS and BDI scales) as well as anxiety (BAI) in female older adults (age 60 yrs. or older) who were stable on pharmacological treatments for MDD and were randomized to the art therapy group (n=31) compared to the control group (n=25). The art therapy intervention consisted of 20, 90-minute sessions that were all provided by the same art therapist and workshops that included a brief relaxation exercise that utilized guided imagery, followed with artistic output and then the participants sharing their thoughts and feelings with the therapist and group members. No adverse events were reported for any of the included RCTs. The strength of the evidence for these studies was rated as low due to small sample sizes and methodological limitations that include unclear information about the randomization process and lack of blinding of patients, clinicians, and outcome assessors.

Table 1. Strength of Evidence for Art Thearpy to Treat MDD

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
Depression	1 SR with 2 RCTs (Meekums, 2015)	DMT (n=54) vs. standard care or WL (n=53) Post-tx	**Change on HAM-D** (mean; 95% CI): **Post-tx:** -7.33, -9.92 to -4.73, p=0.01; NS	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
	1 RCT (Blomdahl, 2018)	PATd + TAU (n=43) vs. TAU alone (n=36) Post-tx	**Change in MADRS-S** (mean; 95% CI): **Post-tx:** 4.00, .38 to 7.63, p=.013; favors PATd +TAU	Yes (-1)	No	No	Yes (-2); small sample size and wide 95% CIs	No	Very low
	1 RCT (Ciasca, 2018)	ART + Med (n=31) vs. Med alone (n=25) Post-tx	**Change in GDS** (mean, [SD]): 3.2(3.4); -0.6(2.32), p=0.007, favors ART **Change in BDI** (mean, [SD]): 8.6(12.8); -1.6(4.86), p=0.025, favors Art therapy	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
Quality of life	1 RCT in Meekums, 2015	DMT (n=10) vs. standard care or WL (n=12)	**Change in MANSA** (mean; 95% CI): 0.30, -0.60	Yes (-1)	No	No	Yes (-2); very small sample size and wide 95% CIs	No	Very low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
		Post-tx	to 1.20, p=0.51; NS						
Anxiety	1 RCT (Ciasca, 2018)	ART + Med (n=31) vs. Med alone (n=25) Post-tx	**Change in BAI** (mean [SD]) -8.9(14.5); 2.9(11.36), p=0.032; favors Art therapy	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
Suicide ideation	1 RCT (Blomdahl, 2018)	PATd+ TAU (n=43) vs. TAU (n=36) Post-tx	**Change in SSI** PATd: 17.1%; TAU: 37.2%; OR: 2.65, 95% CI .87 to 8.05, p=.086; NS	Yes (-1)	No	No	Yes (-2); small sample size and wide 95% CIs	No	Very low

ART: art therapy; BAI: Beck Anxiety Inventory; CI: confidence interval; CT: control group; DMT: Dance Movement Therapy; ES: effective size; f/u: follow-up; mos.: months; GDS: Geriatric Depression Scale; HAM-D: Hamilton Depression Rating Scale; MANSA: Manchester Short Assessment of Quality of Life; MADRS-S: Montgomery-Asberg Depression Rating Scale; NR: not reported; NS: not significant; OR: Odds ratio; PATd: Phenomenological Art Therapy for patients with depression; Post-tx: Posttreatment; QoL: quality of life; RCT: randomized controlled trials; SSI: Scale for Suicide Ideation; SD: standard deviation; TAU: treatment as usual; Tx: treatment; WL: waitlist

Table 2. GRADE Factors Used to Assess the Quality of a Body of Evidence

Evidence Category	Definition
Study Quality (Internal Validity or Risk of Bias)	Study quality considers the overall risk of bias rating of all the studies included in the evidence base. In this review, the overall risk of bias would be the average or median USPSTF rating for studies comprising an evidence base for a key outcome.
Consistency of Evidence	Consistency of evidence refers to the degree of similarity in the direction of effects or the degree of similarity in the effect sizes (magnitude of effect) across individual studies within an evidence base.
Directness of Evidence	Direct evidence directly compares interventions of interest in populations of interest and measures patient-oriented outcomes. Evidence can be indirect if the tested intervention differs from the intervention of interest, the study population differs from the population of

Evidence Category	Definition
	interest, the outcomes differ from those of primary interest, or treatment comparisons have not been tested in head-to-head comparisons.
Precision of Evidence	Precision is the degree of certainty surrounding an estimate of effect with respect to an outcome. Precision is primarily assessed by examining the 95% confidence intervals around the summary effect size.

Link to GRADE Handbook: **http://gdt.guidelinedevelopment.org/app/handbook**

Table 3. Evidence Table for Systematic Review on Art Therapy to Treat MDD

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
Reference: Meekums et al., 2015 **Organization/Country:** Cochrane Library/UK **Purpose:** To examine the effects of DMT for depression with or w/o standard care, compared to no tx or standard care alone, psychological therapies, drug tx, or other physical interventions. Also, to compare the effectiveness of different DMT approaches. **AMSTAR Rating:** High **Overall RoB of Included Studies:** High or unclear (some concerns) due to random sequence generation, allocation concealment, and lack of blinding of patients, treating staff and outcome assessors.	**Databases Searched:** Collaboration Depression, Anxiety and Neurosis Controlled Trials Register (CCDANCTR), CINAHL, World Health Organization's International Clinical Trials Registry Platform (WHO ICTRP), ClinicalTrials.gov, Education Resources Information Center (ERIC) **Dates Searched:** Inception to October 2014. **Inclusion/Exclusion Criteria:** RCTs studying outcomes for people with depression with at least one group being DMT (participatory dance movement with clear psychotherapeutic intent) and facilitated by an individual with a level of training reasonably expected within the country where the trial was conducted **Evidence Base:** 3 RCTs (1 study included adolescents only; 2 studies included adults): only the 2 RCTs that included patients 18 years or older in the meta-analysis will be included in this report.	**Diagnosis:** Depression **Number of Patients:** 147 (107 adults) range per study 31 to 76 **Age (mean yrs):** 40 (for adult pop.) **Gender:** Male (n=51) and female (n=96);	**Intervention:** Dance movement in the presence of a therapist, dance movement interaction w/ a therapist or other group members, or both **Comparators:** WL or standard care **Follow-up:** None reported beyond end of tx score **Outcomes:** Depression levels, drop-out rates, social and occupational functioning, QoL, self-esteem, body image, cost-effectiveness of tx	**DMT vs. standard care or WL** **Depression:** (2 RCTs; n=107); MD: -7.33, 95% CI -9.92 to -4.73, p=0.01, I^2=0% **QoL:** (1 RCT; n=22): MD: 0.30; 95% CI -0.60 to 1.20, p=0.48 No reported AEs

AC: active control; AEs: adverse events; ART: Art Therapy; BL: baseline; CI: confidence interval; ES: effect size; f/u: follow-up; mos.: months; NR: not reported; NS: not significant; RCT: randomized controlled trials; ROB: risk of bias; SD: standard deviation

Table 4. Systematic Review Risk of Bias AMSTAR Checklist Table on Art Therapy for MDD

Question	Meekums et al., 2015
Did the research questions and inclusion criteria for the review include the components of PICO?	Yes
Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	Yes
Did the review authors explain their selection of the study designs for inclusion in the review?	Yes
Did the review authors use a comprehensive literature search strategy?	Yes
Did the review authors perform study selection in duplicate?	Yes
Did the review authors perform data extraction in duplicate?	Yes
Did the review authors provide a list of excluded studies and justify the exclusions?	Yes
Did the review authors describe the included studies in adequate detail?	Yes
Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	Yes
Did the review authors report on the sources of funding for the studies included in the review?	No
If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results? RCTs?	Yes
If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	Yes
Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	Yes
Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	Yes
If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	Yes
Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	Yes
Overall Quality	**High**

RoB: risk of bias

Table 5. AMSTAR Rating of Overall Confidence in Results of the Review

Category	Definition
High	*No or one non-critical weakness*: the systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest.
Moderate	*More than one non-critical weakness*: the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review.
Low or Very Low	*One or more critical flaw(s) with or without non-critical weaknesses*: the systematic review has one or more critical flaws and may not provide an accurate and comprehensive summary of the available studies that address the question of interest.

AMSTAR checklist, go to https://amstar.ca/Amstar_Checklist.php

Table 6. Evidence Table for RCTs on Art Therapy to Treat MDD

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
Reference: Blomdahl et al., 2017 **Purpose:** To examine the effects of manual-based Phenomenological Art Therapy for adult's w/ depression **Setting:** 2 general care clinics and 2 psychiatric outpatient clinics in Sweden **Funding source:** The Health care subcommittee, Region Västra Götaland, VGFOUREG-466011 and Södra Alvsborgs R & D unit, VGFOUSA-340481	**Number of patients:** 79; n=43 PATd + TAU; n=36 TAU alone **Inclusion criteria:** Adults ≥18 years with moderate to severe depression w/o psychotic symptoms. **Exclusion criteria:** Recent traumatic events needing trauma tx, bipolar syndrome, ongoing addiction, psychosis, cognitive disability **Pt. baseline characteristics (PATd + TAU; TAU alone)** **Age (range, %):** 18-25(11.6; 11.1), 26-35(20.9; 30.6), 36-45(32.6; 16.7), 46-55(30.2; 27.8), 56-65(4.7; 13.9) **% male:** 59.4 **% psychiatric comorbidity:** 0(51.2; 52.8); 1(34.9; 44.4); 2(11.6; 2.8); 4(2.3; 0) **Type (n) of psychiatric comorbidity:** PTSD (n=13); mixed anxiety and depression (n=6); generalized anxiety syndrome (n=4); disturbance of activity and attention (n=4)	**Intervention:** PATd was manualized and consisted of 10 1-hr. weekly sessions consisting of various art tasks, serving as a prompt for pt. to paint pictures. Tx carried out by occupational therapists. **Control:** TAU followed ordinary practice and consisted of acupuncture, CBT, ECT, interpersonal therapy, occupational therapy, pharmacological therapy, gemena activity recipe, physiotherapy, psychodynamic therapy, and supportive therapy **Outcomes of interest:** Depression levels (MADRS-S), suicide ideation (SSI), **F/u:** Post-treatment	**Post-Intervention (PATd + TAU; TAU alone)** **Depression:** Pre/post change on MADRS-S (mean btw grp. [SD]): 4.00(.38 to 7.63), p=.013; favors PATd+TAU **Suicide ideation:** PATd: 17.1%; TAU: 37.2%; OR: 2.65, 95% CI .87 to 8.05, p=.086 Adverse events not reported. However, authors did note that some participants did not cope well with answering questions during data collection and dropped out of the trial.	Results suggest that at PATd +TAU statistically significantly improves depression levels compared to TAU alone. There was no significant difference at follow-up between groups in terms of suicide ideation. **Limitations:** Wide CIs for MADRS-S **Study RoB:** Some concerns **Author conflict:** None reported
Reference: Ciasca et al., 2018 **Purpose:** To evaluate if art therapy is effective as an adjunctive tx for depression in the elderly.	**Number of patients:** 56; n=25 in control grp.; n=31 art therapy group **Inclusion criteria:** Lifetime DSM-5 diagnosis of MDD; female gender; age 60 or older; ability to read/write; agreement to take part in the study; stable on	**Intervention:** Pts. participated in 20, 90-min. art therapy sessions all led by same art therapist. Intervention involved 3 grps. of 11 pts. each, however, each pt. was instructed to work on artistic output individually during session.	**Post-treatment GDS*** (mean score for group, mean difference between groups, p-value): Art therapy grp.: -3.2(3.4) (greater reduction in depression)	Results suggest that art therapy statistically significantly reduced severity of depression and anxiety among pts with MDD compared to the CG.

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
Setting: University hospital in Sao Paulo, Brazil **Funding source:** Not reported	pharmacotherapy for depression throughout study. **Exclusion criteria:** Cognitive difficulties suggesting dementia; drug users; pts. w/ degenerative diseases; pts. w/ systematic disorders w/ high morbidity/mortality **Pt. baseline characteristics (Art therapy; Control grp.):** **Age (mean, SD):** 66.1 (5.7); 69.8 (6.4) **First episode after 60 yrs., n (%):** 15(48.4); 14(56) **Number of depressive episodes:** 3.6 (1.7); 3.0 (1.2) **GDS≤5:** 6 (5.3); 9 (3.4) **Medication, n (%):** Antidepressant only: 15(48.4); 14(56) Antidepressant + anxiolytic: 2(6.4); 2(8) Antidepressant + psychotropic: 14(45.2); 9(36)	**Control:** Received no adjuvant tx. **Outcomes of interest:** Depression (measured by the GDS, BDI); anxiety (measured by the BAI) **F/u:** Post-treatment	CG: -0.6(2.32), p=0.007 **Post-treatment BDI*** Art therapy grp.: -8.6(12.8) (greater reduction in depression) CG: -1.6(4.86), p=0.025 **Post-treatment BAI*** Art therapy grp.: -8.9(14.5) (greater reduction in anxiety) CG: -2.9(11.36), p=0.032 *Lower scores on GDS, BDI, and BAI mean less depression or anxiety. Adverse events not reported.	**Limitations:** Integration of brief relaxation and guided imagery to art therapy grp.; attrition; small sample size; different antidepressants used by different pts. **Study RoB:** High **Author conflict:** None reported

AEs: adverse events; BAI: Beck Anxiety Inventory; BDI: Beck Depression Inventory; BL: baseline; CG: control group; CI: confidence interval; ES: effect size; f/u: follow-up; NR: not reported; NS: not significant; OR: odds ratio; RCT: randomized controlled trials; RoB: risk of bias; SD: standard deviation; SMD: standardized mean difference; TAU: treatment as usual;

Table 7. Cochrane Risk of Bias 2.0 Tool for RCTs on Art Therapy to Treat MDD

Reference	Blomdahl, et al. (2018)	Ciasca, et al. (2018)
• Was the allocation sequence generated adequately (e.g., random number table, computer-generated randomization)?	Yes	No
• Was the allocation of treatment adequately concealed (e.g., pharmacy-controlled randomization, concealed envelopes)?	Yes	NI
• Did baseline difference between study groups suggest a problem with randomization?	No	No
Overall ROB for Randomization Process	**Low**	**Some concerns**
Deviation from Intended Intervention (Effect of Assignment)		
• Were participants aware of their assigned intervention during the trial?	Yes	PN
• Were providers and people delivering treatment aware of assigned intervention during trial?	PN	NI
• Were there deviations from the intended intervention that arose because of the experimental context?	PN	NI
• Were these deviations from intended intervention balanced between groups?	NA	NA
• Were these deviations likely to have affected the outcome?	NA	NA
• Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes	NI
Overall ROB of Effect of Assignment	**Low**	**Some concerns**
Missing Outcome Data		
• Were data for this outcome available for all, or nearly all, participants randomized?	Yes	No
• Is there evidence that result was not biased by missing outcome data?	NA	Yes
• Could missingness in the outcome depend on its true value?	NA	NA
• Do the proportions of missing outcome data differ between intervention groups?	NA	NA
• Is it likely that missingness in the outcome depended on its true value?	NA	NA
Overall ROB of Missing Data	**Low**	**Low**
Measurement of the Outcome		
• Was the method of measuring the outcome inappropriate?	No	No
• Could measurement or ascertainment of the outcome have differed between intervention groups?	No	No
• Were outcome assessors aware of the intervention received by study participants?	PY	NI
• Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	NI

Reference	Blomdahl, et al. (2018)	Ciasca, et al. (2018)
• Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	PN
Overall ROB of Measurement of Outcome	**Some concerns**	**Some concerns**
Selection of Reported Results		
• Was the trial analyzed in accordance with a pre-specified plan that was finalized before unblinded outcome data were available for analysis?	NI	NI
Overall ROB of Reported Results	**Some concerns**	**Some concerns**
Overall Study ROB	**Some concerns**	**Some concerns**

*Responses: Y=Yes, PY=Probably Yes, N=No, PN=Probably No, NI=No Information; ROB: risk of bias

Table 8. Cochrane Risk of Bias 2.0 Overall Risk of Bias Judgement

Category	Definition
Low risk of bias	The study is judged to be at **low risk of bias for all domains** for this result.
Some concerns	The study is judged to be at **some concerns** in at least one domain for this result.
High risk of bias	The study is judged to be at **high risk of bias** in at least one domain for this result. OR The study is judged to have **some concerns** for **multiple domains** in a way that substantially lowers confidence in the result.

References

Blomdahl, C., Guregard, S., Rusner, M., & Wijk, H. (2018). Manual-based phenomenological art therapy for individuals diagnosed with moderate to severe depression (PATd): A randomized controlled study. *Psychiatric Rehabilitation Journal, 41*(3), 169-182.

Ciasca, E., Ferreira, R., Santana, C., Forlenza, O., dos Santos, G., Brum, P., & Nunes, P. (2018). Art therapy as an adjuvant treatment for depression in elderly women: A randomized controlled trial. *Brazilian Journal of Psychiatry, 40*, 256-263.

Meekums, B., Karkou, V., & Nelson, E. A. (2015). Dance movement therapy for depression. *Cochrane Database of Systematic Reviews, 2*(CD009895), 1-59.

Tai Chi

Evidence Base

Our searches of the literature identified 3 RCTs that assessed the use of Tai Chi in the treatment of adults with MDD. Yeung et al. (2017) randomized 67 adults with *DSM-IV* MDD receiving no treatment for depression to Tai Chi (n=23), an education program (n=22), or waitlist (n=22). The Tai Chi and education program consisted of 1-hour classes twice a week for 12 weeks. The primary outcomes of interest measured in this study were the response and remission rates for depression.

Yeung et al. (2012) randomized 39 adults with *DSM-IV* MDD to either a 1-hour twice a week for 12-weeks Tai Chi intervention group (n=26) or a waitlist control group (n=13). The primary outcomes of interest measured in this study were the response and remission rates for depression.

Lavretsky et al. (2011) recruited and treated 112 older adults (age ≤ 60) diagnosed with MDD with 10-20 mg per day of escitalopram for 4 weeks. The authors then randomized 73 partial responders who continued to receive escitalopram daily to either Tai Chi Chih (n=36) or a health education program (n=37). The Tai Chi Chih (TCC) sessions were 2-hours once a week and included 10 minutes of warm-up (e.g. stretching and breathing) and 5 minutes of cool-down. The TCC protocol was designed for older adults to address depression, fatigue, and perceived physical limitations and is described by the authors as "meditation through movement." The health education sessions were 2-hours once a week and lasted over the 10-week treatment period using a didactic approach. The sessions consisted of lectures, group discussions, and self-help quizzes to asses participant learning. The purpose of the sessions was to provide education about depression stress, sleep, and other health-related issues that play a role in helping individuals with depression understand and manage their symptoms and factors that contribute to their mood.

Study Quality

Using the Cochrane tool, we rated the RoB of the Yeung (2012) and Lavretsky RCTs as Some Concerns due to lack of information about the randomization process and lack of information about blinding (see **Table 4** for individual quality ratings).

Key Findings

Below, we describe the key findings for the outcomes of interest with the GRADE strength of the evidence (SOE) rating. See **Table 1** for factors that influenced the SOE ratings.

- Evidence from 1 RCT suggests that Tai Chi statistically significantly improves depression levels and remission when compared to waitlist only with the improvements sustained at 24 weeks follow-up (SOE: Low).
- Evidence from 1 RCT suggests that there is no statistically significant difference between Tai Chi and waitlist in improving depression levels or remission rates (SOE: Low).
- Evidence from 1 RCT suggests that Tai Chi as an adjunct to escitalopram statistically significantly improves depression symptom severity but did not reach statistical significance in remission (SOE: Low).

Discussion

The evidence from 1 RCT suggests that Tai Chi compared to waitlist demonstrates positive responses in both depression levels and remission with those improvements being sustained at 24 weeks follow-up. Tai Chi also led to improvements in Clinical Global Impression-Improvement scores with a medium effect size. Another RCT found that Tai Chi as an adjunct treatment to escitalopram is more effective in depression response and remission rates than a health education program as an adjunct to escitalopram. Over time, while both groups improved in terms of depression severity (HAM-D scores), greater reduction was observed among those in the Tai Chi plus escitalopram intervention group. The study also showed that patients who received Tai Chi in addition to escitalopram had better outcomes in health-related quality of life.

However, the findings of a third RCT suggest that, while patients in a Tai Chi intervention group had improved depression response and remission rates compared to those in a waitlist group, the differences were not statistically significant. In addition, there were no statistically significant differences between the Tai Chi and waitlist groups in terms of scores on the HAM-D, Q-LES-Q, CGI-S, or CGI-I.

Table 1. Strength of Evidence for Tai Chi to Treat MDD

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
Depression	3 RCTs (Yeung, 2017; 2012; Lavretsky, 2011)	Tai Chi (23); Education (22); WL (22) 24 wks.	**Response:** OR: 2.26, 95% CI: 0.47-10.84, p>0.05, NS for Tai Chi vs. Edu OR: 2.51, 95% CI: 1.11-5.70, p<0.05, favors Tai Chi over WL **Remission:** OR: 2.40, 95% CI: 0.53-10.85; p>0.05, NS for Tai Chi vs. Edu OR: 2.20, 95% CI: 1.04-4.64, p<0.05, favors Tai Chi over WL	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
		Tai Chi (26); WL (13) 12 wks.	**Change in HAM-D** (mean [SD]): 5.2 (5.1); 4.5 (2.4); Z= -0.23; p=0.82, NS **Remission rate %:** 20; 0, p=0.30, NS	Yes (-1)	No	No	Yes (-1); small sample size	No	Low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
		Tai Chi + esCIT (36); Education + esCIT (37) 14 wks.	**HAM-D (mean, [SD]):** 5.1(3.5); 6.7(4.4) Time (change pre-post tx), p=0.001 Grp. x time, p=0.01, favors Tai Chi **Remission (χ^2)**: 3.68, p<0.06, NS	Yes (-1)	No	No	Yes (-1) small sample size	No	Low
Quality of life	3 RCTs (Yeung, 2017; 2012; Lavretsky, 2011)	Tai Chi (23); Education (22); WL (22) 12 wks.	**SF-36: (mean, [SD]):** Physical function: 778 (18); 718 (184); 703 (174); Pre/Post change ($F_{2,}$): 2.2, p=0.12, NS Emotional: 277 (87); 240 (89); 236 (66) Pre/Post change ($F_{2,}$): 1.0, p=0.38, NS	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
		Tai Chi (26); WL (13) 12 wks.	**Q-LES-Q (mean [SD]):** 0.4 (0.1); 0.4 (0.1); Z= -0.74; p= 0.46, NS	Yes (-1)	No	No	Yes (-1); small sample size	No	Low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
		Tai Chi + esCIT (36); Education + esCIT (37) 14 wks.	**(Mean, (SD):** Physical: 97.3(4.2); 91.1(13.1) Time (change pre-post tx.) (sig.): .97 Grp. x time (sig.): p=0.02, favors Tai Chi Emotional: 83.9(25.2); 71.2(28.3) Time (change pre-post tx) (sig.): p=0.42, NS	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
Anxiety	1 RCT Lavretsky et al., 2011	Tai Chi + esCIT (36); Education + esCIT (37) 14 wks.	**Mean, (SD):** 3.5(2.7); 4.2(3.0) Time (change pre-post tx) (sig.): p=0.001 Grp. x time p=0.27, NS	Yes (-1)	No	No	Yes (-1); small sample size	No	Low

AEs: adverse events; BL: baseline; CI: confidence interval; f/u: follow-up; HAM-A: Hamilton Anxiety Score; HAM-D: Hamilton Depression Score; NR: not reported; NS: not significant; OR: odds ratio, RCT: randomized controlled trials; RoB: risk of bias; SD: standard deviation; SF-36: Short Form Health Survey; Q-LES-Q; Quality of Life Enjoyment and Satisfaction Questionnaire

Table 2. GRADE Factors Used to Assess the Quality of a Body of Evidence

Evidence Category	Definition
Study Quality (Internal Validity or Risk of Bias)	Study quality considers the overall risk of bias rating of all the studies included in the evidence base. In this review, the overall risk of bias would be the average or median USPSTF rating for studies comprising an evidence base for a key outcome.
Consistency of Evidence	Consistency of evidence refers to the degree of similarity in the direction of effects or the degree of similarity in the effect sizes (magnitude of effect) across individual studies within an evidence base.
Directness of Evidence	Direct evidence directly compares interventions of interest in populations of interest and measures patient-oriented outcomes. Evidence can be indirect if the tested intervention differs from the intervention of interest, the study population differs from the population of interest, the outcomes differ from those of primary interest, or treatment comparisons have not been tested in head-to-head comparisons.
Precision of Evidence	Precision is the degree of certainty surrounding an estimate of effect with respect to an outcome. Precision is primarily assessed by examining the 95% confidence intervals around the summary effect size.

Link to GRADE Handbook: **http://gdt.guidelinedevelopment.org/app/handbook**

Table 3. Evidence Table for RCTs on Tai Chi to Treat MDD

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
Reference: Yeung et al. 2017 **Purpose:** To assess the effectiveness of Tai Chi as monotherapy for MDD **Setting:** NR **Funding source:** National Center for Complementary and Integrative Health, Bethesda, MD	**Number of patients:** 67; n=23 Tai Chi; n=22 Education; n=22 Waitlist **Inclusion criteria:** Male and female adults between 18 and 70 years with a diagnosis of Major Depressive Disorder according to the *DSM-IV* who self-identified as being of Chinese ethnicity and fluent in Mandarin or Cantonese **Exclusion criteria:** Pts with primary psychiatric disorder other than MDD; history of psychosis, mania, or severe cluster B personality disorder; unstable medical conditions; suicidal or self-injurious; regular practice of Tai Chi or other mind-body interventions in last 12 mos.; current or planned use of potentially confounding txs during study including antidepressants, psychotherapy, or CAM, other mind-body interventions **Pt. baseline characteristics (Tai Chi; Education; WL):** **Age (mean yrs., SD):** 53 (14); 55 (9); 55 (15) **Gender (% female):** 74%; 73%; 68%	**Intervention:** Tai Chi, 1-hr. twice a wk. for 12 wks. **Control:** Education (didactic training, discussion of stress, mental health, depression), 1-hr. twice a wk. for 12 wks.; WL, contacted for assessment at wks. 6, 12, 18, 24 **Outcomes of Interest:** Depression levels as measured by HDRS (positive response defined as decrease in total scores of 50% or more; remission defined as HDRS ≤ 7); QoL as measured by SF-36; global improvement and severity as measured by CGI-I and CGI-S respectively; AEs **Follow-up:** 24 weeks	**<u>End of Tx (12 wks.)</u>** **Depression:** Tai Chi vs. Edu.: OR: 8.90, 95% CI: 1.17-67.70, p<.05; favors Tai Chi Tai Chi vs. WL: OR: 2.11, 95% CI: 1.01-4.46, p<.05; favors Tai Chi **Remission:** Tai Chi vs. Edu.: OR: 4.40, 95% CI: 0.78-24.17; NS Tai Chi vs. WL: OR: 3.01, 95% CI: 1.25-7.10, p<.05; favors Tai Chi HDRS (mean, [SD]): Tai Chi: 9 (5); Edu.: 15 (7); WL: 134 (7) Pre/Post change ($F_{2,}$): 2.9, p=.07 BDI (mean, [SD]): Tai Chi: 11 (6); Edu: 18 (9); WL: 19 (10) Pre/Post change ($F_{2,}$): 2.7, p=.08 **Global severity/improvement:** CGI-S: (mean, [SD]): Tai Chi: 3(1); Edu: 3(1); WL: 3(1) Pre/Post change ($F_{2,}$): 2.0, p=.15	**Conclusion:** The findings suggest there was a statistically significant positive response in the Tai Chi group in depression levels and remission when compared to waitlist only with the improvements being sustained at 24-wk. f/u. Tai Chi led to improvements in CGI-I scores with a medium effect size. There was no statistically significant difference between the Tai Chi, education, or waitlist grps. in terms of HDRS or BDI scores. **Limitations:** Small sample size; multiple statistical analyses comparing continuous outcome measurements may have led to false positives or type 1 errors; considerable drop-out rate (22% after randomization); lack of blinding; results may not be generalizable. **Study RoB:** High due to lack of information about randomization process, lack of information about blinding, and high attrition. **Author conflict:** Yes, Dr. Wayne is the founder/owner of Tree of Life Tai Chi Center; Dr. Denninger holds a position at the Benson-Henry Institute for Mind Body Medicine at MGH, which is paid by pts. and insurers for running relaxation/mindfulness clinical programs and markets merchandise.

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
			CGI-I: (mean, [SD]): Tai Chi: 2(1); Edu: 3(1); WL: 3(1) Pre/post change ($F_{2,}$): 3.4, p=.04; Tai Chi vs. WL ES: 0.78 **Quality of life:** SF-36: (mean, [SD]): Physical function: Tai Chi: 778 (18); Edu: 718 (184); WL: 703 (174) Pre/Post change ($F_{2,}$): 2.2, p=.12 Emotional well-being: Tai Chi: 277 (87); Edu: 240 (89); WL: 236 (66) Pre/Post change ($F_{2,}$): 1.0, p=.38 Pain: Tai Chi: 151(33); Edu: 127(32); 145(34) Pre/Post change ($F_{2,}$): 1.8, p=.18; **F/u (24 wks.)** **Depression:** Tai Chi vs. Edu.: OR: 2.26, 95% CI: 0.47-10.84; NS Tai Chi vs. WL: OR: 2.51, 95% CI: 1.11-5.70, p<.05; favors Tai Chi **Remission:** Tai Chi vs. Edu: OR: 2.40, 95% CI: 0.53-10.85; NS	

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
			Tai Chi vs. WL: OR: 2.20, 95% CI: 1.04-4.64, p<.05; favors Tai Chi **AEs**: None reported	
Reference: Yeung et al. 2012 **Purpose:** To assess the possible efficacy feasibility, and safety of using Tai Chi for treatment MDD **Setting:** Not reported **Funding source:** Not reported	**Number of patients:** 39; n=26 Tai Chi; n=13 WL **Inclusion criteria:** Self-identified as being of Chinese ethnicity and fluent in Mandarin and/or Cantonese; 18-70 yrs. of age; *DSM-IV* diagnosis of MDD; baseline score of ≤12 on the HAM-D **Exclusion criteria:** Primary psychiatric diagnosis other than MDD; history of psychosis, mania, severe cluster B personality disorder; unstable medical conditions; current active suicidal or self-injurious potential; regular Tai Chi or other forms of mind-body practice in past 3 mos. **Pt. baseline characteristics (Tai Chi; WL):** **Age (mean yrs., SD):** 54 (12); 58 (7) **Gender (% male):** 23%; 23%	**Intervention:** Tai Chi, 1-hr. twice/wk. for 12 wks. **Control:** Waitlist **Outcomes of Interest:** Depression and remission as measured by HAM-D (positive response defined as decrease of ≤50%, remission defined as score of ≥7); QoL as measured by the Q-LES-Q-SF **Follow-up:** 12 weeks	**<u>End of Tx - 12 weeks (Tai Chi; WL)</u>** **Depression** Response rate (%): 24%; 0%; p=0.15, NS Remission rate (%): 20%; 0%; p=0.30, NS Change in HAM-D (mean [SD]): 5.2 (5.1); 4.5 (2.4); Z= -0.23; p=0.82, NS **Global severity/improvement** Change in CGI-S (mean [SD]): 1.0 (1.0); 0.67 (1.2); Z= -0.74; p= 0.50, NS Change in CGI-I (mean [SD]): 3.0 (1.2); 3.5 (1.0); Z= -1.4; p= 0.21, NS **Quality of life** Change in Q-LES-Q (mean [SD]): 0.4 (0.1); 0.4 (0.1); Z= -0.74; p= 0.46, NS **AEs**: None reported	**Conclusion:** The findings suggest that, while pts. in the Tai Chi group had improved depression response and remission rates compared to those in the waitlist group, the differences were not statistically significant. There were no statistically significant differences between Tai Chi and waitlist groups in terms of scores on the HAM-D, Q-LES-Q, CGI-S, or CGI-I. **Limitations:** Small sample size; unclear whether pt. improvement resulted from Tai Chi training or attention/social support accompanying the intervention; results may not be generalizable to other populations as pts. were predominately Chinese immigrants **Study RoB:** Some concern **Author conflict:** None reported

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
Reference: Lavretsky et al. 2011 **Purpose:** To assess the possible efficacy of the mind-body exercise, Tai Chi Chih as an adjunct to escitalopram in the tx of geriatric depression **Setting:** NR **Funding source:** NR	**Number of patients:** 73; n=36 TCC + escitalopram; n=37 health education + escitalopram **Inclusion criteria:** Current episode of MDD, HDRS score of ≤16 at baseline, MMSE score of ≤26 **Exclusion criteria:** History of other psychiatric illness or substance use disorder/dependence, severe medical illness, acute suicidal or violent behavior, any other central nervous system disease or dementia, unable to participate in TCC due to mobility issues **Pt. baseline characteristics (TCC + esCIT; HE + esCIT):** **Age (mean yrs., SD):** 69.1 (7.0); 72.0 (7.4) **Gender (% female):** 64%; 60%	**Intervention:** TCC + esCIT, 2 hrs/wk. + 10-20 mg/d **Control:** HE + esCIT, 2 hrs./wk. + 10-20 mg/d **Outcomes of Interest:** Depression levels and remission as measured by HDRS, anxiety as measured by Hamilton Anxiety Scale, QoL as measured by the Medical Outcomes Study 36-item Short Form Health Survey **Follow-up:** 14 weeks	**End of Tx (10 wks.)** **Depression (mean, [SD]):** Tai Chi + esCIT: 5.1(3.5) HE + esCIT: 6.7(4.4) Time (change pre-post tx) (sig.): .001 Grp. x time (sig.): .01; favors Tai Chi + esCIT **Anxiety (mean, [SD]:** Tai Chi + esCIT: 3.5(2.7) HE + esCIT: 4.2(3.0) Time (change pre-post tx) (sig.): .001 Grp. x time (sig.): .27; favors Tai Chi + esCIT **QoL (mean, [SD]:** Physical: Tai Chi + esCIT: 97.3(4.2) HE + esCIT: 91.1(13.1) Time (change pre-post tx) (sig.): .97 Grp. x time (sig.): .02 Emotional: Tai Chi + esCIT: 83.9(25.2) HE + esCIT: 71.2(28.3) Time (change pre-post tx) (sig.): .42 Grp. x time (sig.): .003 **AEs:** None reported	**Conclusions:** Depression response and remission rates were higher among pts. in the Tai Chi group compared to the health education group. Over time, while both groups improved in terms of depression severity (HAM-D scores), greater reduction was observed among those in the Tai Chi intervention group. Greater improvements in health-related quality of life were seen among the Tai Chi group when compared to the health education group. **Limitations:** small sample size, short f/u **Study RoB:** Some concern **Author conflict:** None reported

AEs: adverse events; BL: baseline; CI: confidence interval; esCIT: escitalopram; f/u: follow-up; HAM-A: Hamilton Anxiety Score; HAM-D: Hamilton Depression Score; HE: health education; NR: not reported; NS: not significant; OR: odds ratio, RCT: randomized controlled trials; RoB: risk of bias; SD: standard deviation; SF-36: Short Form Health Survey; TC: Tai Chi; TCC: Tai Chi Chih; Q-LES-Q; Quality of Life Enjoyment and Satisfaction Questionnaire; WL: Waitlist

Table 4. Cochrane Risk of Bias 2.0 Tool for RCTs Tai Chi for MDD

Reference	Yeung et al. 2017	Yeung et al. 2012	Lavretsky et al., 2011
• Was the allocation sequence generated adequately (e.g., random number table, computer-generated randomization)?	Yes	Yes	Yes
• Was the allocation of treatment adequately concealed (e.g., pharmacy-controlled randomization, concealed envelopes)?	Yes	Yes	Yes
• Did baseline difference between study groups suggest a problem with randomization?	No	No	No
Overall RoB for Randomization Process	**Low**	**Low**	**Low**
Deviation from Intended Intervention (Effect of Assignment)			
• Were participants aware of their assigned intervention during the trial?	Yes	Yes	Yes
• Were providers and people delivering treatment aware of assigned intervention during trial?	PY	Yes	NI
• Were there deviations from the intended intervention that arose because of the experimental context?	PY	PN	PN
• Were these deviations from intended intervention balanced between groups?	Yes	NA	NA
• Were these deviations likely to have affected the outcome?	NA	NA	NA
• Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes	Yes	Yes
Overall RoB of Effect of Assignment	**Some concerns**	**Low**	**Low**
Missing Outcome Data			
• Were data for this outcome available for all, or nearly all, participants randomized?	No	Yes	Yes
• Is there evidence that result was not biased by missing outcome data?	PN	NA	NA
• Could missingness in the outcome depend on its true value?	PN	NA	NA
• Do the proportions of missing outcome data differ between intervention groups?	NA	NA	NA
• Is it likely that missingness in the outcome depended on its true value?	NA	NA	NA
Overall RoB of Missing Data	**Low**	**Low**	**Low**
Measurement of the Outcome			
• Was the method of measuring the outcome inappropriate?	No	No	No
• Could measurement or ascertainment of the outcome have differed between intervention groups?	No	No	No
• Were outcome assessors aware of the intervention received by study participants?	NI	No	No

Reference	Yeung et al. 2017	Yeung et al. 2012	Lavretsky et al., 2011
• Could assessment of the outcome have been influenced by knowledge of intervention received?	Yes	NA	NA
• Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	NA	NA
Overall RoB of Measurement of Outcome	**High**	**Low**	**Low**
Selection of Reported Results			
• Was the trial analyzed in accordance with a pre-specified plan that was finalized before unblinded outcome data were available for analysis?	Yes	NI	NI
Overall RoB of Reported Results	**Low**	**Some concerns**	**Some concerns**
Overall Study RoB	**High**	**Some concerns**	**Some concerns**

*Responses: Y=Yes; PY=Probably Yes; N=No; PN=Probably No; NA=Not Applicable; NI=No Information; RoB: risk of bias

Table 5. Cochrane Risk of Bias 2.0 Overall Risk of Bias Judgement

Category	Definition
Low risk of bias	The study is judged to be at **low risk of bias for all domains** for this result.
Some concerns	The study is judged to be at **some concerns** in at least one domain for this result.
High risk of bias	The study is judged to be at **high risk of bias** in at least one domain for this result. OR The study is judged to have **some concerns** for **multiple domains** in a way that substantially lowers confidence in the result.

References

Lavretsky, H., Altstein, L., Olmstead, R., Ercoli, L., Riparetti-Brown, M., …Irwin, M. (2011). Complementary use of Tai Chi Chih augments escitalopram treatment of geriatric depression: A randomized controlled trial. *American Journal of Geriatric Psychiatry, 19*(10), 839-850.

Yeung, A., Feng, R., Kim, D., Wayne, P., Yeh, G., Baer, L., …Fava, M. (2017). A pilot, randomized controlled study of Tai Chi with passive and active controls in the treatment of depressed Chinese Americans. *Journal of Clinical Psychiatry, 78*(5), e522-e528.

Yeung, A., Lepoutre, V., Wayne, P., Yeh, G., Slipp, L., Fava, M., …Benson, H. (2012). Tai Chi treatment for depression in Chinese Americans. *American Journal of Physical Medicine & Rehabilitation, 91*(10), 863-870.

Music Therapy

Evidence Base

Our searches of the literature identified 1 RCT that met inclusion criteria and assessed the effects of music therapy (MT) in the treatment of adults diagnosed with depression. See **Table 3** for details about the patients, interventions, outcomes and findings of the identified study.

Erkkila et al. (2011) looked at the possible efficacy of MT as an adjunct treatment to standard care for depression compared with standard care alone in adults aged 18-50 years. In this study, 79 patients diagnosed with unipolar depression (*ICD-10-CM* or *DSM-III-R*) were randomized to receive either MT or treatment as usual (TAU). Standard care consisted of 5 to 6 individual psychotherapy sessions, antidepressants, and psychiatric counseling. Music therapy consisted of 20 bi-weekly sessions which were 60 minutes each. Patients were permitted to continue taking medication during the study. The main purpose of the MT intervention was to have patients engage in expressive musical interaction with a therapist supporting and facilitating the therapeutic process using musical elements (i.e. rhythm, harmony, melody, etc.). The primary outcome of interest was depression severity. Secondary outcomes include anxiety, general functioning, and quality of life.

Study Quality

Using the Cochrane tool, we rated the RoB of the Erkkila (2011) RCT as having some concerns due to lack of blinding. (see **Table 4** for individual quality ratings).

Key Findings

Below, we describe the key findings for the outcomes of interest with the GRADE strength of the evidence (SOE) rating. See **Table 1** for factors that influenced the SOE ratings.

- Evidence from 1 RCT suggests that, while music therapy in conjunction with standard care statistically significantly improves depression upon completion of treatment (3 mos.) compared to standard care alone (SOE: Low), those differences do not reach statistical significance at 6 months follow-up (SOE: Very low).

- Evidence from 1 RCT suggests that, while music therapy in conjunction with standard care statistically significantly improves anxiety upon completion of treatment (3 mos.) compared to standard care alone (SOE: Low), those differences do not reach statistical significance at 6 months follow-up (SOE: Very low).

- Evidence from 1 RCT suggests that, while music therapy in conjunction with standard care statistically significantly improves functioning upon completion of treatment (3 mos.) compared to standard care alone (SOE: Low), those differences do not reach statistical significance at 6 months follow-up (SOE: Very low).

- Evidence from 1 RCT suggests there is no statistical significance between music therapy in conjunction with standard care and standard care alone in improving health-related quality of life (SOE: Very low).

Discussion

Overall, the results of the Erkkila et al. RCT suggests that music therapy in conjunction with standard care demonstrates positive responses in depression levels, anxiety, and functioning when compared to standard care immediately following treatment, which in this case, was 3 months. These effects were statistically significant with effect sizes in the medium-to-large range (0.65 for depression and 0.49 for anxiety). However, these differences did not reach statistical significance in the 6-month follow-up assessments. The overall strength of the evidence for all the reported outcomes of interest was rated low to very low (**See Table 1**). This is largely due to limitations in the methodological quality of the study and the small sample size.

Table 1. Strength of Evidence for Music Thearpy to Treat MDD

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
Depression	1 RCT (Erkkila, 2011)	MT (n=33) vs. standard care (n=46) 3 mos.	**3 mos. (mean [SD]**, 95% CI): 14.10 (8.77); 16.43 (9.33), 0.59 to 8.70, p=0.03; favors MT Effect size, d = 0.65	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
		MT (n=33) vs. standard care (n=46) 6 mos.	**6 mos. f/u (mean [SD]):** 14.48 (9.60); 14.74 (10.65), -1.05 to 7.94, p=0.13; NS	Yes (-1)	No	No	Yes (-2); small sample size and wide 95% CIs	No	Very low
Anxiety	1 RCT (Erkkila, 2011)	MT (n=33) vs. standard care (n=46) 3 mos.	**3 mos. (mean [SD]**, 95% CI): 7.37 (3.99); 8.00 (4.11), 0.09 to 3.55, p=0.04; favors MT Effect size, d = 0.49	Yes (-1)	No	No	Yes (-1); small sample size	No	Low

		MT (n=33) vs. standard care (n=46) 6 mos.	**6 mos. f/u (mean [SD]):** 7.21 (4.15); 7.29 (4.75), -0.38 to 3.67, p=0.11; NS	Yes (-1)	No	No	Yes (-2); small sample size and wide 95% CIs	No	Very low
Functioning	1 RCT (Erkkila, 2011)	MT (n=33) vs. standard care (n=46) 3 mos.	**3 mos. (mean [SD]**, 95% CI): 70.00 (9.37); 66.78 (9.61), -8.93 to -0.24, p=0.04; favors MT Effect size, d = 0.62	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
		MT (n=33) vs. standard care (n=46) 6 mos.	**6 mos. f/u:** (mean, [SD]; 95% CI): 72.90 (13.89); 70.74 (12.64), -10.48 to 1.35, p=0.13; NS	Yes (-1)	No	No	Yes (-2); small sample size and wide 95% CIs	No	Very low
Quality of Life	1 RCT (Erkkila, 2011)	MT (n=33) vs. standard care (n=46) 3 mos.	**3 mos. (mean [SD]**, 95% CI): 66.70 (20.10); 62.59 (18.20), -11.40 to 2.40, p=0.20; NS	Yes (-1)	No	No	Yes (-2); small sample size and wide 95% CIs	No	Very low
		MT (n=33) vs. standard care (n=46) 6 mos.	**6 mos. (mean [SD]**, 95% CI): 67.93 (18.51); 64.60 (18.74), -11.83 to 3.57, p=0.29; NS	Yes (-1)	No	No	Yes (-2); small sample size and wide 95% CIs	No	Very low

CI: confidence interval; CT: control group; ES: effective size; f/u: follow-up; mos.: months; MT: music therapy; NR: not reported; NS: not significant; RCT: randomized controlled trials; SD: standard deviation

Table 2. GRADE Factors Used to Assess the Quality of a Body of Evidence

Evidence Category	Definition
Study Quality (Internal Validity or Risk of Bias)	Study quality considers the overall risk of bias rating of all the studies included in the evidence base. In this review, the overall risk of bias would be the average or median USPSTF rating for studies comprising an evidence base for a key outcome.
Consistency of Evidence	Consistency of evidence refers to the degree of similarity in the direction of effects or the degree of similarity in the effect sizes (magnitude of effect) across individual studies within an evidence base.
Directness of Evidence	Direct evidence directly compares interventions of interest in populations of interest and measures patient-oriented outcomes. Evidence can be indirect if the tested intervention differs from the intervention of interest, the study population differs from the population of interest, the outcomes differ from those of primary interest, or treatment comparisons have not been tested in head-to-head comparisons.
Precision of Evidence	Precision is the degree of certainty surrounding an estimate of effect with respect to an outcome. Precision is primarily assessed by examining the 95% confidence intervals around the summary effect size.

Link to GRADE Handbook: **http://gdt.guidelinedevelopment.org/app/handbook**

Table 3. Evidence Table for RCTs on Music Therapy to Treat MDD

Study Details	Study Population	Treatment	Results	Conclusion/ Limitations
Reference: Erkkila et al., 2011 Purpose: To examine the efficacy of music therapy added to standard care compared w/ standard care alone in adults with depression Setting: Music Therapy Clinic for Research and Training, Univ. of Jyvaskyla, Finland Funding source: NEST (New and Emerging Science and Technology) program of the European Commission; Centres of Excellence in research, Academy of Finland	Number of patients: 79; n=33 MT; n=46 TAU alone Inclusion criteria: Adults w/ unipolar depression between the age of 18 and 50. Exclusion criteria: History of repeated suicidal behavior, psychosis; acute substance misuse; severity of depression prevented pts. from participating in measurements/engaging in verbal conversation; unable to speak/understand Finnish Pt. baseline characteristics (MT; TAU alone) Age (mean [SD]): 35.8 (9.0); 35.5 (10.5) % female: 75.8; 80.4 % current medication (self-reported): Any antidepressant: 66.7; 76.1 SSRI: 48.5; 43.5 SNRI: 15.2; 20.0	Intervention: MT was delivered by music therapists in 20 bi-weekly sessions lasting 60 min. each. Musical expression in the sessions was based on a restricted selection of music instruments. Pt. and therapist used identical instrumentation. Sessions were recorded for processing/discussion. Control: TAU included short-term psychotherapy (5 to 6 individual sessions) conducted by nurses trained in depression, antidepressants, and psychiatric counselling. Outcomes of interest: Depression levels (MADRS), anxiety (HADS-A), functioning (GAF), health-related QoL (RAND-36) F/u: 6 mos.	<u>End of tx (3 mos. f/u): (MT; TAU alone)</u> Depression (mean [SD], 95% CI): 14.10 (8.77); 16.43 (9.33), 0.59 to 8.70, p=0.03; favors MT Effect size, d = 0.65 Anxiety (mean [SD], 95% CI): 7.37 (3.99); 8.00 (4.11), 0.09 to 3.55, p=0.04; favors MT Effect size, d = 0.49 Functioning (mean [SD], 95% CI): 70.00 (9.37); 66.78 (9.61), -8.93 to -0.24, p=0.04; favors MT Effect size, d = 0.62 Health-related QoL (mean [SD], 95% CI): 66.70 (20.10); 62.59 (18.20), -11.40 to 2.40, p=0.20; NS <u>6 mos. f/u: (MT; TAU alone)</u> Depression (mean [SD], 95% CI): 14.48 (9.60); 14.74 (10.65), -1.05 to 7.94, p=0.13; NS Anxiety (mean [SD], 95% CI): 7.21 (4.15);	Results suggest that MT added to TAU statistically significantly improves depression levels, anxiety, and functioning compared to TAU alone upon tx completion. These differences were no longer statistically significant at 6 mos. f/u. Limitations: Sample size large enough to detect effect in primary outcome post-tx, but not at 6 mos. f/u. Study RoB: Some concern Author conflict: None reported

Study Details	Study Population	Treatment	Results	Conclusion/ Limitations
			7.29 (4.75), -0.38 to 3.67, p=0.11; NS Functioning (mean [SD], 95% CI): 72.90 (13.89); 70.74 (12.64), -10.48 to 1.35, p=0.13; NS Health-related QoL (mean [SD], 95% CI): 67.93 (18.51); 64.60 (18.74), -11.83 to 3.57, p=0.29; NS Adverse events were reported. Two pts. (one in each arm) experienced significant worsening of their depression, leading them to quit the study early. One pt. in the control grp. developed severe low back pain.	

AEs: adverse events; BDI: Beck Depression Inventory; BL: baseline; CG: control group; CI: confidence interval; ES: effect size; f/u: follow-up; MT: music therapy; NR: not reported; NS: not significant; OR: odds ratio; QoL: quality of life; RCT: randomized controlled trials; RoB: risk of bias; SD: standard deviation; SF-36: Short-Form 36; SMD: standardized mean difference; TAU: treatment as usual

Table 4. Cochrane Risk of Bias 2.0 Tool for RCTs on Music Therapy to Treat MDD

Reference	Erkkila, et al. (2011)
• Was the allocation sequence generated adequately (e.g., random number table, computer-generated randomization)?	Yes
• Was the allocation of treatment adequately concealed (e.g., pharmacy-controlled randomization, concealed envelopes)?	Yes
• Did baseline difference between study groups suggest a problem with randomization?	No
Overall ROB for Randomization Process	**Low**
Deviation from Intended Intervention (Effect of Assignment)	
• Were participants aware of their assigned intervention during the trial?	PY
• Were providers and people delivering treatment aware of assigned intervention during trial?	PY
• Were there deviations from the intended intervention that arose because of the experimental context?	PN
• Were these deviations from intended intervention balanced between groups?	NA
• Were these deviations likely to have affected the outcome?	NA
• Was an appropriate analysis used to estimate the effect of assignment to intervention?	NA
Overall ROB of Effect of Assignment	**Low**
Missing Outcome Data	
• Were data for this outcome available for all, or nearly all, participants randomized?	Yes
• Is there evidence that result was not biased by missing outcome data?	NA
• Could missingness in the outcome depend on its true value?	NA
• Do the proportions of missing outcome data differ between intervention groups?	NA
• Is it likely that missingness in the outcome depended on its true value?	NA
Overall ROB of Missing Data	**Low**
Measurement of the Outcome	
• Was the method of measuring the outcome inappropriate?	No
• Could measurement or ascertainment of the outcome have differed between intervention groups?	PN
• Were outcome assessors aware of the intervention received by study participants?	Yes (while outcome assessor was masked, they unintentionally became aware of some participants' status)
• Could assessment of the outcome have been influenced by knowledge of intervention received?	Yes

Reference	Erkkila, et al. (2011)
• Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN
Overall ROB of Measurement of Outcome	**Some concern**
Selection of Reported Results	
• Was the trial analyzed in accordance with a pre-specified plan that was finalized before unblinded outcome data were available for analysis?	Yes
Overall ROB of Reported Results	**Low**
Overall Study ROB	**Some concern**

*Responses: Y=Yes, PY=Probably Yes, N=No, PN=Probably No, NI=No Information; ROB: risk of bias

Table 5. Cochrane Risk of Bias 2.0 Overall Risk of Bias Judgement

Category	Definition
Low risk of bias	The study is judged to be at **low risk of bias for all domains** for this result.
Some concerns	The study is judged to be at **some concerns** in at least one domain for this result.
High risk of bias	The study is judged to be at **high risk of bias** in at least one domain for this result. OR The study is judged to have **some concerns** for **multiple domains** in a way that substantially lowers confidence in the result.

References

Erkkila, J., Punkanen, M., Fachner, J., Ala-Ruona, E., Pontio, I., Tervaniemi, M., …Gold, C. (2011). Individual music therapy for depression: Randomised controlled trial. *The British Journal of Psychiatry, 199*, 132-139.

Yoga

Evidence Base

Our searches of the literature identified 1 SR that assessed the use of yoga in the treatment of adults diagnosed with depression. See **Table 3** for details about the patients, interventions, outcomes and findings of the identified studies.

In brief, Vollbehr et al. (2018) conducted a SR with meta-analysis that evaluated the effects of hatha yoga for adults diagnosed with depression on depressive symptoms. The evidence base for the SR included a total of 11 RCTs enrolling 1,327 patients (range per study 20 to 620) that were used in the meta-analysis.

Study Quality

Using the AMSTAR instrument, we rated the quality of the Vollbehr review as low as there was no evidence that a protocol was developed prior to conduct of the review (See **Table 4** for the review ratings). The authors of this review assessed the RoB of the RCTs using the Clinical Trial Assessment Measure (CTAM). The overall RoB of the trials included in the Vollbehr review was mainly low to moderate due to lack of adequate description of the intervention, lack of reporting around attrition, and lack of blinding of patients, providers, and outcome assessors.

Key Findings

Below, we describe the key findings for the outcomes of interest with the GRADE strength of the evidence (SOE) rating. See **Table 1** for factors that influenced the SOE ratings.

- Evidence from 6 RCTs suggests that yoga statistically significantly improves depressive symptoms in adults diagnosed with clinical depression when compared to a psychoeducation control (SOE: Very low).
- Evidence from 6 RCTs suggest that there is no significant difference between yoga and an active control (other than psychoeducation) in improving depressive symptoms in adults diagnosed with clinical depression (SOE: Very low).
- Evidence from 6 RCTs suggest there is no significant difference between yoga and TAU in improving depressive symptoms in adults diagnosed with clinical depression (SOE: Very low).
- Evidence from 4 RCTs suggest there is no significant difference between yoga and an active control (psychoeducation and other) in improving depressive symptoms in adults diagnosed with clinical depression at ≥6 months follow-up (SOE: Very low).

Discussion

Overall, the findings of the Vollbehr review suggest that yoga compared to treatment as usual or compared to all active controls, show no significant effect on symptoms of depression. However, when comparing yoga to psychoeducation control, yoga did lead to reductions in symptoms of depression. At six months follow-up or longer however, yoga showed no significant effect when compared to active control.

The overall strength of the evidence for yoga was very low and limited due to limitations in the methodological quality of the RCTs (e.g. lack of blinding, attrition), statistical imprecision, and considerable heterogeneity. Larger, more rigorous studies are needed to fully assess the effectiveness of yoga in the treatment of depression.

Table 1. Strength of Evidence for Yoga to Treat MDD

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
Depression	1 SR with 11 RCTs (Vollbehr, 2019)	Yoga (n=392*) Control includes active control or TAU (n=784*) Post-tx *1 RCT (Field, 2012) did not specify how many pts. were randomized into yoga	**Yoga vs. active control (all, psychoeducation and other) or TAU:** (11 RCTs; n=1,209); SMD: -0.13, 95% CI -0.49 to 0.22, p=0.47, I^2=77%; NS **Yoga vs. psychoeducation control:** (6 RCTs; n=283); SMD: -0.52, 95% CI -0.96 to -0.08, p=0.02, I^2=56%; favors yoga **Yoga vs. other active control:** (6 RCTs; n=978); SMD: 0.28, 95% CI -0.07 to 0.63, p=0.12, I^2=65%; NS	Yes (-1)	Yes (-1); considerable heterogeneity	No	Yes (-1); wide 95% CIs	No	Very low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
			Yoga vs. TAU: (6 RCTs; n=912); SMD: -0.64, 95% CI -1.41 to 0.13, p=0.10, I^2=93%; NS						
		Yoga (n=153) Control includes active control (n=573) ≥6 mos. f/u	**Yoga vs. active control (all, psychoeducation, and other):** (4 RCTs; n=1,209); SMD: -0.14, 95% CI -0.60 to 0.33, p=0.56, I^2=78%; NS						

CI: confidence interval; CT: control group; DASS-21: Depression, Anxiety, and Stress Scale; ES: effective size; f/u: follow-up; mos.: months; NR: not reported; NS: not significant; RCT: randomized controlled trials; SD: standard deviation; SF-12: Short Form Health Survey; TAU: treatment as usual; QoL: quality of life

Table 2. GRADE Factors Used to Assess the Quality of a Body of Evidence

Evidence Category	Definition
Study Quality (Internal Validity or Risk of Bias)	Study quality considers the overall risk of bias rating of all the studies included in the evidence base. In this review, the overall risk of bias would be the average or median USPSTF rating for studies comprising an evidence base for a key outcome.
Consistency of Evidence	Consistency of evidence refers to the degree of similarity in the direction of effects or the degree of similarity in the effect sizes (magnitude of effect) across individual studies within an evidence base.
Directness of Evidence	Direct evidence directly compares interventions of interest in populations of interest and measures patient-oriented outcomes. Evidence can be indirect if the tested intervention differs from the intervention of interest, the study population differs from the population of interest, the outcomes differ from those of primary interest, or treatment comparisons have not been tested in head-to-head comparisons.
Precision of Evidence	Precision is the degree of certainty surrounding an estimate of effect with respect to an outcome. Precision is primarily assessed by examining the 95% confidence intervals around the summary effect size.

Link to GRADE Handbook: **http://gdt.guidelinedevelopment.org/app/handbook**

Table 3. Evidence Table for Systematic Review on Yoga to Treat MDD

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
Reference: Vollbehr et al., 2018 **Organization/Country:** University of Groningen/Netherlands **Purpose:** To examine the effects of hatha yoga in treating acute, chronic and/or treatment-resistant depression **AMSTAR Rating:** Low **Overall RoB of Included Studies:** Low to High due to random sequence generation, allocation concealment, and lack of blinding of patients, treating staff and outcome assessors.	**Databases Searched:** Medline, Cochrane Library, Current Controlled Trials, Clinical Trials.gov, NHR Centre for Reviews and Dissemination, PsycINFO and CINAHL **Dates Searched:** Inception to June 2018. **Inclusion/Exclusion Criteria:** RCTs comparing a yoga intervention to WL, TAU, or active control; included majority of pts. with MDD (diagnosed according to DSM or ICD criteria); adults aged 18-65 yrs.; English language; included continuous measurement of improvement or dichotomous measure of remission of mood and/or anxiety symptoms at pre- and post-intervention using self-report or clinician-rated scales. Excluded MBSR and MBCT excluded due to focus on meditation. **Evidence Base:** 18 RCTs in SR; 13 RCTs used in meta-analysis and reported on in this report.	**Diagnosis:** Depression **Number of Patients:** 1,327, range per study 20 to 620 **Age (range):** 18-65 **Gender:** Majority of pts. were female; 6 studies were female only; 7 studies were mixed gender	**Intervention:** 13 examined yoga ranging in duration from 5 to 13 wks. In a few of the studies (k=3), yoga practice occurred 1x75min./wk. For the remaining studies (k=10), yoga practice ranged from 1 to 5 times per week for 20 to 120min. In most studies (k=8), home practice was encouraged. **Comparators:** Active interventions (psychoeducation, healthy living classes, mindfulness, walking) and TAU **Follow-up:** 4 RCTs had a follow-up period of 6 months or more. **Outcomes:** Depression levels	**Post-treatment:** **Yoga vs. active control (all, psychoeducation, and other) or TAU** **Depression:** (11 RCTs; n=1,209); SMD: -0.13, 95% CI -0.49 to 0.22, p=0.47, I^2=77%; NS **Yoga vs. psychoeducation control** **Depression:** (6 RCTs; n=283); SMD: -0.52, 95% CI -0.96 to -0.08, p=0.02, I^2=56%; favors yoga **Yoga vs. other active control** **Depression:** (6 RCTs; n=978); SMD: 0.28, 95% CI -0.07 to 0.63, p=0.12, I^2=65%; NS **Yoga vs. TAU** **Depression:** (6 RCTs; n=912); SMD: -0.64, 95% CI -1.41 to 0.13, p=0.10, I^2=93%; NS **≥6 mos. f/u:** **Yoga vs. active control (all, psychoeducation, and other)** **Depression:** (4 RCTs; n=1,209); SMD: -

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
				0.14, 95% CI -0.60 to 0.33, p=0.56, I^2=78%; NS No reported AEs

AC: active control; AEs: adverse events; BL: baseline; CI: confidence interval; ES: effect size; f/u: follow-up; mos.: months; NR: not reported; NS: not significant; RCT: randomized controlled trials; ROB: risk of bias; SD: standard deviation; TAU: treatment as usual

Table 4. Systematic Review Risk of Bias AMSTAR Checklist Table on Yoga for MDD

Question	Vollbehr et al., 2018
Did the research questions and inclusion criteria for the review include the components of PICO?	Yes
Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	No
Did the review authors explain their selection of the study designs for inclusion in the review?	Yes
Did the review authors use a comprehensive literature search strategy?	Yes
Did the review authors perform study selection in duplicate?	Yes
Did the review authors perform data extraction in duplicate?	Yes
Did the review authors provide a list of excluded studies and justify the exclusions?	Yes
Did the review authors describe the included studies in adequate detail?	Yes
Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	Yes
Did the review authors report on the sources of funding for the studies included in the review?	No
If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results? RCTs?	Yes
If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	Yes
Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	Yes
Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	Yes
If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	Yes
Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	Yes
Overall Quality	**Low**

RoB: risk of bias

Table 5. AMSTAR Rating of Overall Confidence in Results of the Review

Category	Definition
High	*No or one non-critical weakness*: the systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest.
Moderate	*More than one non-critical weakness*: the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review.
Low or Very Low	*One or more critical flaw(s) with or without non-critical weaknesses*: the systematic review has one or more critical flaws and may not provide an accurate and comprehensive summary of the available studies that address the question of interest.

AMSTAR checklist, go to https://amstar.ca/Amstar_Checklist.php

References

Vollbehr, N., Bartels-Velthuis, A., Nauta, M., Castelein, S., Steenhuis, L., Hoenders, H., & Ostafin, B. (2018). Hatha yoga for acute, chronic and/or treatment-resistant mood and anxiety disorders: A systematic review and meta-analysis. *PLoSONE, 13*(10), 1-28.

Meditation

Evidence Base

Our searches of the literature identified 3 RCTs that assessed the use of meditation in the treatment of adults with MDD. Ionson et al. (2018) randomized 83 older adults (aged 60-85) with a current major depressive episode (confirmed with Mini Neuropsychiatric Interview) and receiving treatment as usual (TAU) to either a Sahaj Samadhi meditation (SSM) group (n=40) or a TAU alone group (n=43). The SSM groups consisted of 4 or more participants led by certified teachers. An instructional phase consisted of 4 sessions with a duration of 90-120 min. each where participants received a sound (mantra) and a method of using the mantra. The correct practice and understanding of meditation was then reinforced. After the instructional phase was complete, 60-minute weekly sessions occurred over the next 11 weeks. Participants were expected to attend at least 75% of the sessions and engage in home SSM for 20 min. two times a day. Participants in the SSM group continued to receive TAU. The TAU control group continued to receive TAU including non-structured supportive therapy and/or antidepressant medications for 12 weeks. The outcomes of interest measured in this study were changes and severity of depressive symptoms, anxiety, quality of life, and adverse effects.

Sharma et al. (2017) randomized 25 adults with *DSM-IV-TR* MDD who had not responded to >8 weeks of antidepressants to either an adjunct Sudarshan Kriya yoga (SKY), a breathing-based meditation intervention group (n=13) or waitlist control group (n=12). The SKY sessions were provided for 1 hour bi-weekly over 12 weeks and consisted of two phases. The first phase (week 1) was a 6-session group program in which participants engaged in sitting meditation, yoga poses, and stress education. The second phase (weeks 2-8) consisted of weekly 1.5 hr. SKY sessions. Participants were also asked to practice SKY at home for 20-25 minutes per day. Those in the waitlist group were offered the SKY intervention once the study was completed. The primary outcomes of interest measured in this study include symptoms of depression and anxiety.

Winnebeck et al. (2017) recruited and randomized 74 adults diagnosed with *DSM-IV-TR* MDD to either a brief mindfulness-based meditation intervention (n=38) or a psychoeducation and resting group control group (n=36). The psychoeducation and resting control group learned about the signs and causes of depression, specifically the role of stress and the need to balance stress by taking time to rest, which patients in this group were instructed to do in order to disengage from negative thinking. Participants in the intervention group engaged 25 min. of guided meditation twice daily as well as informal short practices including breathing spaces (time to pause, relax, and decide what to do next). The primary outcome of interest measured in this study is severity of depressive symptoms.

Study Quality

Using the Cochrane tool, we rated the RoB of the Ionson (2018), Sharma (2017), and Winnebeck (2017) RCTs as Some Concerns due to lack of information about the randomization process and lack of information about blinding (see **Table 4** for individual quality ratings).

Key Findings

Below, we describe the key findings for the outcomes of interest with the GRADE strength of the evidence (SOE) rating. See **Table 1** for factors that influenced the SOE ratings.

- Evidence from 3 RCTs suggest that meditation alone or as an adjunct to treatment as usual (including antidepressants) or waitlist statistically significantly improves depressive symptoms when compared to treatment as usual alone, medication alone, or active control at 8-14 weeks follow-up (SOE: Low).
- Evidence from 2 RCTs suggest that meditation as an adjunct to treatment as usual (including antidepressants) statistically significantly improves anxiety symptoms 8-12 weeks follow-up (SOE: Low).
- Evidence from 1 RCT suggests that there is no significant difference between meditation as an adjunct to treatment as usual compared to treatment as usual alone in improving quality of life (SOE: Very low).
- Evidence from 1 RCT suggests that meditation in conjunction with treatment as usual statistically significantly improves remission compared to treatment as usual alone (SOE: Low)

Discussion

Overall, the findings from three RCTs suggest that meditation-based interventions offered as an adjunctive therapy to treatment as usual (including antidepressants) or waitlist reduce symptoms and severity of depression. The findings of the Ionson (2018) RCT also found that meditation used in conjunction with treatment as usual is associated with greater improvements in anxiety symptoms compared to treatment as usual alone. The Ionson study also looked at the outcome of quality of life, however, no significant difference was found between meditation and treatment as usual alone.

The overall strength of the evidence for meditation-based interventions ranged from low to very low. In general, the strength of the evidence was limited due to limitations in the methodological quality of the RCTs (e.g. lack of blinding, unclear randomization process), small sample sizes, and very short follow-up periods. Larger, more rigorously designed studies with longer follow-up periods are needed.

Table 1. Strength of Evidence for Meditation to Treat MDD

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
Depression	3 RCTs (Ionson, 2018; Sharma, 2017; Winnebeck, 2017)	Meditation + TAU (40); TAU (43); 12 wks.	**HRSD-17 (mean point change, [95% CI]):** SSM vs. TAU.: -3.96 (-6.00 to -1.91), p=0.0009; -1.30 (-2.65 to 0.05), p=0.060; Difference: -2.66 (-5.05 to -0.26), p=0.030; favors SSM	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
		Meditation + Meds. (13); WL + Meds. (12) 8 wks.	**HRSD-17 (mean point change, [95% CI]):** SSM vs. TAU.: -10.27 (-5.04 to -15.50), p=.0032, favors SKY **BDI (mean point change, [95% CI]):** SSM vs. TAU.: -15.48 (-8.34 to -22.62),	Yes (-1)	No	No	Yes (-1); small sample size	No	Low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
			p=.0101, favors SKY						
		Meditation (38); Psychoeducation + resting (36) 14 wks.	**BDI-II (pre-post change [SE], p):** MBI: -17.41 (1.50), p=0.000 Edu + resting: -9.30 (1.59), p=0.000; favors MBI	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
Remission	1 RCT (Ionson 2018)	Meditation + TAU (40); TAU (43); 12 wks.	**HRSD (OR, [95% CI]):** 3.36 (1.06 to 10.64), p=0.040, favors MED	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
Quality of life	1 RCT (Ionson, 2018)	Meditation + TAU (40); TAU (43); 12 wks.	**QOLPSV (mean point change, [95% CI]):** SSM vs. TAU.: 10.23 (2.95 to 17.50), p=.007; 8.31 (0.63 to 16.00), p=.035; Difference: 1.91 (-8.54 to 12.37), p=0.72; NS	Yes (-1)	No	No	Yes (-2); single study with a small sample and wide 95% CI	No	Very low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
Anxiety	2 RCTs (Ionson, 2018; Sharma, 2017)	Meditation + TAU (40); TAU (43); 12 wks.	**GAI mean, (SD):** 3.5(2.7); 4.2(3.0) Time (change pre-post tx) (sig.): p=0.001 Grp. x time p=0.27, NS	Yes (-1)	No	No	Yes (-1); small sample size	No	Low
		Meditation + Meds. (13); WL + Meds. (12) 8 wks.	**BAI (mean point change, [95% CI]):** SSM vs. TAU.: -5.19 (-0.93 to -9.34), p=.0097, favors SKY	Yes (-1)	No	No	Yes (-1); small sample size	No	Low

AEs: adverse events; BAI: Beck Anxiety Inventory; BDI: Beck Depression Inventory; BL: baseline; CI: confidence interval; f/u: follow-up; GAI: Geriatric Anxiety Inventory; HRSD: Hamilton Depression Rating Score; NR: not reported; NS: not significant; OR: odds ratio, QOLPSV: Quality of Life Profile Senior Version; RCT: randomized controlled trials; RoB: risk of bias; SD: standard deviation

Table 2. GRADE Factors Used to Assess the Quality of a Body of Evidence

Evidence Category	Definition
Study Quality (Internal Validity or Risk of Bias)	Study quality considers the overall risk of bias rating of all the studies included in the evidence base. In this review, the overall risk of bias would be the average or median USPSTF rating for studies comprising an evidence base for a key outcome.
Consistency of Evidence	Consistency of evidence refers to the degree of similarity in the direction of effects or the degree of similarity in the effect sizes (magnitude of effect) across individual studies within an evidence base.
Directness of Evidence	Direct evidence directly compares interventions of interest in populations of interest and measures patient-oriented outcomes. Evidence can be indirect if the tested intervention differs from the intervention of interest, the study population differs from the population of interest, the outcomes differ from those of primary interest, or treatment comparisons have not been tested in head-to-head comparisons.
Precision of Evidence	Precision is the degree of certainty surrounding an estimate of effect with respect to an outcome. Precision is primarily assessed by examining the 95% confidence intervals around the summary effect size.

Link to GRADE Handbook: **http://gdt.guidelinedevelopment.org/app/handbook**

Table 3. Evidence Table for RCTs on Meditation to Treat MDD

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
Reference: Ionson et al. 2018 **Purpose:** To assess the effectiveness of Sahaj Samadhi meditation (SSM) in the treatment of heart rate variability and depression **Setting:** NR **Funding source:** Innovation Fund of the Alternative Funding plan of the Academic Health Sciences Centres of Ontario, London, Ontario, Canada and the Schulich	**Number of patients:** 83; n=48 SSM; n=47 TAU **Inclusion criteria:** Male and female adults between 60 and 85 w/ a current major depressive episode (MINI); stable physical health; no severe cardiac episode in past 12 mos.; able to sit comfortably for 30-45 min. w/o major pain/discomfort; able to hear/follow verbal instructions w/ eyes closed; willing/able to attend 4 initial SSM training sessions and 75% follow-up appts. **Exclusion criteria:** Pts participating in other studies or regularly practiced any other type of meditation; diagnosed w/ stroke, transient ischemic attack, heart disease, or seizure in past 6 mos.; score	**Intervention:** SSM for 4 90-120 min. sessions consecutively, then 60 min./wk. for 11 wks. + home practice for 20 min. twice daily. **Control:** TAU for 12 wks. Including antidepressant meds. and/or non-structured supportive therapy. **Outcomes of Interest:** Depression levels as measured by HRSD; improvement and	**End of Tx (12 wks.)** **Depression (mean point change, [95% CI]):** SSM vs. TAU.: -3.96 (-6.00 to -1.91), p=0.0009; -1.30 (-2.65 to 0.05), p=0.060; Difference: -2.66 (-5.05 to -0.26), p=0.030; favors SSM	**Conclusion:** The findings suggest there was a statistically significant positive response in the SSM group in depression levels, anxiety levels, and clinical global improvement when compared to TAU alone. There was no statistically significant difference between the SSM or TAU alone grps. in terms of QoL scores. **Limitations:** No active comparator to control variables; small sample size **Study RoB:** Some concerns due to lack of information about blinding. **Author conflict:** Yes, one author is the Director of Research and Health Promotion for the Art of Living

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
Research Opportunity Project, London, Ontario, Canada	<24 on MMSE; experienced suicidal thoughts; had other significant mental health diagnoses. **Pt. baseline characteristics (SSM; TAU):** **Age (mean yrs., SD):** 69.45 (5.8); 68.30 (6.5) **Gender (% female):** 60%; 77% **Taking antidepressant med. (%):** 75%; 60% **Taking ≥2 antidepressant meds. (%):** 37%; 35% **SSRIs (%):** 33%; 37% **SNRIs (%):** 18%; 21% **Mirtazapine/bupropion (%):** 38%; 23% **Tricyclic antidepressants (%):** 5%; 7% **Antipsychotic use (%):** 20%; 7% **Serotonin antagonist and reuptake inhibitor (%):** 4%; 9%	severity as measured by CGI-I; anxiety as measured by GAI; QoL as measured by QOLPSV; AEs **Follow-up:** 12 weeks	**Anxiety (mean point change, [95% CI]):** SSM vs. TAU.: -3.71 (-5.15 to -2.26), p<.0001; -1.34 (-2.77 to 0.10), p=0.07; Difference: -2.37 (-4.37 to -0.36), p=0.021; favors SSM **Clinical Global Improvement (mean point change, [95% CI]):** SSM vs. TAU.: 2.85 (2.3 to 3.4), p<.0001; 3.65 (3.26 to 4.04), p<.0001; Difference: -0.80 (-1.46 to -0.15), p=0.18; favors SSM **QoL (mean point change, [95% CI]):** SSM vs. TAU.: 10.23 (2.95 to 17.50), p=.007; 8.31 (0.63 to 16.00), p=.035; Difference: 1.91 (-8.54 to 12.37), p=0.72; NS **AEs**: None reported	Foundation, Canada, and supervised staff providing SSM; one author has received research funding from Satellite Healthcare for a mindfulness meditation trial.
Reference: Sharma et al. 2017	**Number of patients:** 25; n=13 SKY; n=12 WL (delayed yoga)	**Intervention:** SKY, (week 1) 3.5 hrs./day;	**<u>End of Tx - 8 weeks (SKY; WL)</u>**	**Conclusion:** The findings suggest pts. w/ inadequate response to

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
Purpose: To assess the possible efficacy feasibility, and safety of using SKY as an adjunct tx. for MDD **Setting:** Univ. of PA Presbyterian Hospital Clinic and Translational Research Center (CTRC) **Funding source:** American Psychiatric Association and Indo-American Psychiatric Association, NIH (grant number: UL1TR000003)	**Inclusion criteria:** Adults aged 18-67 yrs. w/ diagnosis of MDD according to DSM-IV-TR criteria; stable on ≥8 wks. of antidepressant; HDRS-17 score of ≥14 at baseline. **Exclusion criteria:** Bipolar disorder, psychosis, substance abuse, ADHD, pregnancy, epilepsy; initiating psychotherapy and/or other yoga/meditation programs. **Pt. baseline characteristics (SKY; WL):** **Age (mean yrs., SD):** 39.4 (13.9); 34.8 (13.6) **Gender (% female):** 69.2%; 75%	(Weeks 2-8) 1.5 hrs./wk. + 20-25 min./day at home. **Control:** Waitlist (delayed yoga) **Outcomes of Interest:** Depression as measured with HDRS-17 and BDI; anxiety as measured with BAI **Follow-up:** 8 weeks	**Depression (mean point change, [95% CI]):** SSM vs. TAU.: -10.27 (-5.04 to -15.50), p=.0032, favors SKY **Self-report depression (mean point change, [95% CI]):** SSM vs. TAU.: -15.48 (-8.34 to -22.62), p=.0101, favors SKY **Anxiety (mean point change, [95% CI]):** SSM vs. TAU.: -5.19 (-0.93 to -9.34), p=.0097, favors SKY **AEs**: None reported	antidepressants may benefit from SKY as an adjunct tx. Patients in SKY statistically significantly improved in depression and anxiety symptom outcomes when compared to those in the waitlist group **Limitations:** Small sample size; lack of active comparator grp., lack of blinding. **Study RoB:** Some concerns due to lack of information about blinding and allocation concealment. **Author conflict:** One author received grants from AHRQ, Alkermes, Forest, NIMH, Otsuka, PharmaNeuroboost, and Roche, and has acted as advisor/consultant for Alkemes, AstraZeneca, Bristol-Myers Squibb, Cerecor, Eli Lilly, Forest, Gerson Lehman Group, GlaxoSmithKline, Guidepoint Global, Lundbeck, MedAvante, Merck, Neuronetics, Novartis, Ortho-McNeil, Otsuka, Pamlab, Pfizer, Shire, Sunovion and Takeda
Reference: Winneback et al. 2017 **Purpose:** To assess the possible efficacy of a brief mindfulness meditation intervention in reducing depressive symptoms. **Setting:** NR	**Number of patients:** 74; n=38 MBI; n=36 psychoeducation + resting **Inclusion criteria:** Adults age 25-60 w/ current diagnosis of MDD; chronic or recurrent lifetime history of depression w/ onset before age 19 and either chronic persistence of symptoms or a history of at least 3 previous episodes of depression within last 2 yrs.; self-reported clinical severity of current symptoms according to BDI-II; fluent in German.	**Intervention:** MBI, 25 min., twice per day formal meditation **Control:** EDU + resting, 25 min., twice per day **Outcomes of Interest:** Depression symptom severity as measured by BDI-II	**<u>End of Tx (2 wks.)</u>** **Depression (pre-post change [SE], p):** MBI: -17.41 (1.50), p=0.000 Edu + resting: -9.30 (1.59), p=0.000; favors MBI	**Conclusions:** Brief mindfulness meditation training statistically significantly improved self-reported depressive symptoms when compared to a psychoeducation + resting control. **Limitations:** Very short follow-up, reliance on self-reported outcomes. **Study RoB:** Some concerns due to lack of information around blinding. **Author conflict:** None reported

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
Funding source: German Research Foundation (grant: BA2255 2-1)	**Exclusion criteria:** History of psychosis or mania, current eating disorder, OCD, current self-harm, current substance abuse/dependence; history of TBI; current CBT tx. **Pt. baseline characteristics (MBI; EDU + resting):** **Age (mean yrs., SD):** 42.3 (12.4); 40.7 (12.2) **Gender (% female):** 61%; 59%	**Follow-up:** 2 weeks	**AEs:** None reported	

AEs: adverse events; BL: baseline; CI: confidence interval; f/u: follow-up; EDU: education; GAI: Geriatric Anxiety Inventory; HRSD: Hamilton Depression Rating Score; MBI: mind body intervention; NR: not reported; NS: not significant; QOLPSV: Quality of Life Profile Senior Version; RCT: randomized controlled trials; RoB: risk of bias; SD: standard deviation; wks.: weeks

Table 4. Cochrane Risk of Bias 2.0 Tool for RCTs Meditation for MDD

Reference	Ionson et al. 2018	Sharma et al. 2017	Winnebeck et al. 2017
• Was the allocation sequence generated adequately (e.g., random number table, computer-generated randomization)?	Yes	Yes	Yes
• Was the allocation of treatment adequately concealed (e.g., pharmacy-controlled randomization, concealed envelopes)?	Yes	NI	Yes
• Did baseline difference between study groups suggest a problem with randomization?	No	No	No
Overall RoB for Randomization Process	**Low**	**Some concerns**	**Low**
Deviation from Intended Intervention (Effect of Assignment)			
• Were participants aware of their assigned intervention during the trial?	NI	Yes	NI
• Were providers and people delivering treatment aware of assigned intervention during trial?	No	NI	NI
• Were there deviations from the intended intervention that arose because of the experimental context?	No	No	No
• Were these deviations from intended intervention balanced between groups?	NA	NA	NA
• Were these deviations likely to have affected the outcome?	NA	NA	NA
• Was an appropriate analysis used to estimate the effect of assignment to intervention?	No	Yes	No
Overall RoB of Effect of Assignment	**Some concerns**	**Low**	**Some concerns**
Missing Outcome Data			
• Were data for this outcome available for all, or nearly all, participants randomized?	No	Yes	Yes
• Is there evidence that result was not biased by missing outcome data?	Yes	NA	NA
• Could missingness in the outcome depend on its true value?	NA	NA	NA
• Do the proportions of missing outcome data differ between intervention groups?	NA	NA	NA
• Is it likely that missingness in the outcome depended on its true value?	NA	NA	NA
Overall RoB of Missing Data	**Low**	**Low**	**Low**
Measurement of the Outcome			
• Was the method of measuring the outcome inappropriate?	No	No	No
• Could measurement or ascertainment of the outcome have differed between intervention groups?	No	No	No
• Were outcome assessors aware of the intervention received by study participants?	No	No	No

Reference	Ionson et al. 2018	Sharma et al. 2017	Winnebeck et al. 2017
• Could assessment of the outcome have been influenced by knowledge of intervention received?	NA	NA	NA
• Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	NA	NA
Overall RoB of Measurement of Outcome	**Low**	**Low**	**Low**
Selection of Reported Results			
• Was the trial analyzed in accordance with a pre-specified plan that was finalized before unblinded outcome data were available for analysis?	Yes	Yes	Yes
Overall RoB of Reported Results	**Low**	**Low**	**Low**
Overall Study RoB	**Some concerns**	**Some concerns**	**Some concerns**

*Responses: Y=Yes; PY=Probably Yes; N=No; PN=Probably No; NA=Not Applicable; NI=No Information; RoB: risk of bias

Table 5. Cochrane Risk of Bias 2.0 Overall Risk of Bias Judgement

Category	Definition
Low risk of bias	The study is judged to be at **low risk of bias for all domains** for this result.
Some concerns	The study is judged to be at **some concerns** in at least one domain for this result.
High risk of bias	The study is judged to be at **high risk of bias** in at least one domain for this result. OR The study is judged to have **some concerns** for **multiple domains** in a way that substantially lowers confidence in the result.

References

Ionson, E., Limbachia, J., Rej, S., Puka, K., Newman, R., Wetmore, S., …Vasudev, A. (2018). Effects of Sahaj Samadhi meditation on heart rate variability and depressive symptoms in patients with late-life depression. *The British Journal of Psychiatry*, 1-7.

Sharma, A., Barrett, M., Cucchiara, A., Gooneratne, N., & Thase, M. (2017). A breathing-based meditation intervention for patients with major depressive disorder following inadequate response to antidepressants: A randomized pilot study. *Journal of Clinical Psychiatry, 78*(1), e59-e63.

Winnebeck, E., Fissler, M., Gartner, M., Chadwick, P., & Barnhofer, T. (2017). Brief training in mindfulness meditation reduces symptoms in patients with a chronic or recurrent lifetime history of depression: A randomized controlled study. *Behaviour Research and Therapy, 99*, 124-130.

Transcranial Magnetic Stimulation (TMS) or Repetitive (r)TMS

Evidence Base

Our searches of the literature identified 2 systematic reviews (SRs) and 1 subsequently published RCT that evaluated the efficacy of rTMS on adults with treatment resistant depression (TRD). The first SR by Brunoni et al. (2017) evaluated the efficacy and acceptability of the different modalities of rTMS used for MDD by performing a network meta-analysis (Brunoni et al., 2017). For this report, we did not consider evidence from this review of studies on newer techniques of rTMS, such as synchronized rTMS, pulsed rTMS, deep rTMS, or rTMS with priming stimulation. We only considered the evidence from studies of more conventional forms of rTMS that includes high or low frequency rTMS applied to either the left or right dorsolateral prefrontal cortex (DLPC) or applied bilaterally. The overall evidence base for the Brunoni review included 81 RCTs enrolling 4,233 patients, most of whom (74%) were diagnosed with TRD. In most of the studies included in this review, rTMS was an add-on or augmentative therapy to pharmacotherapy (88%) and was compared to sham rTMS with or without pharmacotherapy.

The second review was conducted by Health Quality Ontario (Ontario's Health Technology Assessment center, 2016). This review focused specifically on the efficacy of high frequency (≥5 Hz) rTMS applied to the left DLPC. Because this review had more direct evidence on this application of rTMS, we did not include evidence from the Brunoni review on high frequency left DLPC. The evidence base for this review included 23 RCTs comparing rTMS to sham and enrolling a total of 1,156 patients with TRD. In 16 studies, patients received rTMS while receiving antidepressants, and in seven studies patients did not receive any antidepressant during rTMS treatment. The RCT by Yesavage et al., (2018), published subsequent to the 2 SRs randomized 164 adults with TRD to receive rTMS applied to the left DLPC (n=81) or to receive sham rTMS (n=83). This study enrolled only US military veterans receiving care at a VA medical centers.

Our searches also identified an RCT that considered the use of rTMS among patients with first episode major depression. Wang et al (2017) randomized 43 patients with first episode depression to receive rTMS combined with an antidepressant drug (paroxetine, n=22) or to sham rTMS plus an antidepressant (n=21).

The primary outcomes in all the studies considered as evidence for rTMS were response to therapy (typically defined as 50% or greater improvement from baseline according to the study's primary depression scale), any improvement in symptoms of depression, complete remission of depression, and adverse events. Two studies included in the Brunoni review (Blumberg et al. 2012 & Blumberg et al. 2016) and the RCT by Yesavage also reported on suicidality. See **Tables 3 and 6** for more information about the studies included as evidence for the use of rTMS to treat depression.

Study Quality

Using the AMSTAR instrument, we rated the quality of the review by Brunoni as moderate primarily because this review did not include a list of excluded studies with reasons for exclusion (see **Table 4** for quality ratings of SRs). The authors of the review used the Cochrane risk of bias (ROB) tool to rate the methodological quality of the included studies. According to the authors, 21.0%, 67.9%, and 11.1% of studies had an overall low, unclear, and high ROB. The unclear ratings were due mostly to lack of reporting of randomization or allocation procedures and/or imperfect blinding. We rated the quality of the review by Health Quality Ontario as High quality. The authors of this review also used the Cochrane tool

to rate individual study ROB. Most of the included studies received a quality rating of moderate as only three studies performed allocation concealment and not all studies blinded patients.
Using the Cochrane tool, we rated the ROB of the two additional RCTs as Low. In both studies, patients and outcomes assessors were adequately blinded and there was little to no attrition (See **Table 7** for quality ratings of individual studies).

Key Findings

Below, we describe the key findings for the outcomes of interest with the GRADE strength of the evidence (SOE) rating. See **Table 1** for factors that influenced the SOE ratings.

Treatment-Resistant Depression

High-Frequency rTMS (>5 Hz)

- Evidence from 17 RCTs suggests that high frequency rTMS is more effective than sham rTMS in reducing symptoms of depression. (SOE: Moderate)
- Evidence from 13 RCTs suggests that high frequency rTMS leads to significantly higher rates of remission from depression compared to sham rTMS. (SOE: Moderate)
- Evidence from 20 RCTS suggests that significantly more patients with treatment-resistant depression responded to high frequency rTMS compared to sham rTMS. (SOE: Moderate)
- Evidence from 1 RCT suggests that there is no difference between high frequency rTMS and sham rTMS in reducing symptoms of suicidality or PTSD. (SOE: Moderate)
- Evidence from 1 RCT suggests that there is no difference between high frequency rTMS and sham rTMS in improving mental or physical quality of life. (SOE: Moderate)

Low-Frequency rTMS (<5 Hz)

- Evidence from 1 SR (with 81 studies overall) suggests that significantly more patients with treatment resistant depression respond to low frequency rTMS compared to sham rTMS. (SOE: Moderate)
- Evidence from 2 RCTs suggests that there is no difference between low frequency rTMS and sham rTMS in resolving suicide ideation. (SOE: Low)

Bi-lateral rTMS (any frequency)

- Evidence from 1 SR (with 81 studies overall) suggests that significantly more patients with treatment resistant depression respond to bilateral rTMS compared to sham rTMS. (SOE: Moderate)
- Evidence from 1 SR (with 81 studies overall) suggests that bilateral rTMS leads to significantly higher rates of remission from depression compared to sham rTMS. (SOE: Moderate)
- Evidence from 2 RCTs suggests that patients who receive bilateral rTMS are more likely to experience resolution of suicidal ideation compared to patients who receive sham rTMS. (SOE: Moderate)

Acute Depression

- Evidence from 1 RCT suggests that rTMS plus paroxetine reduces symptoms of depression among adults with acute depression compared to sham rTMS plus paroxetine. (SOE: Low)
- Evidence from 1 RCT suggests that significantly more patients who received rTMS plus paroxetine responded to and experienced remission from depression compared to patients who received sham rTMS and paroxetine. (SOE: Low)

Discussion

Overall, the findings of the RCTs that made up the evidence base for rTMS suggest that significantly more patients with treatment resistant depression who received active rTMS alone or as an adjunct to medication respond to therapy, experience improvement in overall symptoms of depression, and achieve remission from depression compared to patients who received sham rTMS. The strength of the evidence supporting the findings for rTMS for these outcomes was rated as moderate due to some concerns about the methodological quality of some of the included RCTs. Evidence from two RCTs further suggests that active rTMS is more effective in reducing symptoms of suicidality among patients with treatment resistant depression compared to sham rTMS. However, no differences were found between active and sham rTMS for mental or physical quality of life. This finding, however, is based on evidence from one relatively small RCT.

A single RCT found that significantly more patients with first episode (or acute) depression who received rTMS plus paroxetine responded to therapy, experienced improvement in overall symptoms of depression and achieved remission from depression compared to patients who received sham rTMS plus paroxetine. The strength of the evidence supporting these findings was low due to the very small sample size of the study.

Few studies reported on adverse events. Among those that did, headache and scalp discomfort were the most commonly reported events. In general, lack of reporting of adverse events was a limitation of the evidence supporting the use of rTMS for treating depression. Additional limitations included small evidence base for patients with first time depression, and limited length of treatment and follow-up to assess adverse events associated with repeated exposures to rTMS or for follow-up periods longer than 6 months.

Table 1. Strength of Evidence for Transcranial Magnetic Stimulation (TMS) to Treat MDD

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
High Frequency (≥5 Hz) dorsolateral rTMS vs. Sham for Treatment Resistant Depression									
Depression symptoms	17 RCTs in 1 SR + 1 RCT Health Quality Ontario, 2016 and Yesavage et al 2018	HF rTMS (683) vs Sham (637) 2 weeks to 6 mos	**In SR:** WMD: 2.31 points, 95% CI 1.19–3.43, p< 0.001, favors rTMS, I^2=19.8% **Yesavage:** **Posttx:** MD: 1.28, 95% CI -1.42 to 3.97, p=0.34, NS **24 wks:** MD: 0.62, 95% CI -2.59 to 3.94, p=0.68, NS	Yes (-1)	No	No	No	No	Moderate
Remission	13 RCTs in 1 SR + 1 RCT Health Quality Ontario, 2016 and Yesavage et al 2018	HF rTMS (683) vs Sham (637) 2 weeks to 6 mos	**In SR:** RR: 2.20, 95% CI 1.44–3.38, p< 0.001, favors rTMS, I^2=0.0% **Yesavage:** Posttx: OR: 1.16, 95% CI 0.59 to 2.26, p=0.67, NS 24 wks: OR: 1.55, 95% CI 0.62 to 3.86, p=0.35, NS	Yes (-1)	No	No	No	No	Moderate

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
Response	20 RCTs in 1 SR Health Quality Ontario, 2016	HF rTMS (602) vs Sham (554) 2 weeks to 6 mos	RR: 1.72, 95% CI 1.13–2.62, p=0.011, favors rTMS, I^2=46.4%	Yes (-1)	No	No	No	No	Moderate
Suicidality (BSI)	1 RCT Yesavage et al 2018	HF rTMS (81) vs Sham (83) 2 weeks to 6 mos	**Posttx:** MD: 0.08, 95% CI -1.46 to 1.62, p=0.91, NS **24 wks:** -MD: 0.54, 95% CI -2.25 to 1.17, p=0.53, NS	No	No	No	Yes (-1), wide confidence interval	No	Moderate
PTSD Symptom Severity (CAPS)	1 RCT Yesavage et al 2018	HF rTMS (81) vs Sham (83) 2 weeks to 6 mos	**Posttx:** MD: 5.20, 95% CI -0.49 to 10.89, p=0.07, NS **24 wks:** MD: 4.47, 95% CI -0.69 to 9.64, p=0.09, NS	No	No	No	Yes (-1), wide confidence interval	No	Moderate
Quality of life (Veterans RAND-36	1 RCT Yesavage et al 2018	HF rTMS (81) vs Sham (83) 2 weeks to 6 mos	**Posttx** (Physical): MD: -1.32, 95% CI -3.61 to 0.97, p=0.27, NS (Mental):MD: -1.76, 95% CI -5.91 to 2.39), p=0.40, NS	No	No	No	Yes (-1), wide confidence interval	No	Moderate

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
			<u>24 wks</u> (Physical): MD: 0.08, 95% CI -2.67 to 2.83, p=0.96, NS (Mental): MD: -0.12, 95% CI -4.48 to 4.24, p=0.96, NS						
Low Frequency (≤1 Hz) rTMS vs. Sham for Treatment Resistant Depression									
Response	81 RCTs included in 1 SR Brunoni et al., 2017 Note: does not indicate how many studies contributed to specific comparison	LF rTMS vs Sham Overall patients=4,233 F/u: NR	OR: 2.48, 95% CI 2.33 to 4.61, favors rTMS	Yes (-1)	No	No	No	No	Moderate
Resolution of Suicide Ideation	2 RCTs included in 1 SR Brunoni et al., 2017 Note: does not indicate how many studies	LF rTMS vs Sham n=4,233 patients overall (number of patients in each group NR	OR: 1.59, 95% CI 0.61 to 4.12, p=0.33, NS	Yes (-1)	No	No	Yes (-1); wide confidence interval	No	Low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
	contributed to specific comparison	F/u: NR							
Bi-lateral rTMS (any frequency) vs Sham for Treatment Resistant Depression									
Response	81 RCTs included in 1 SR Brunoni et al., 2017 Note: does not indicate how many studies contributed to specific comparison	Bi-lateral rTMS vs Sham n=4,233 (number of patients in each group NR F/u: NR	OR: 3.39, 95% CI, 1.91 to 6.02, favors bilateral rTMS	Yes (-1)	No	No	No	No	Moderate
Resolution of Suicide Ideation	2 RCTs included in 1 SR Brunoni et al., 2017	Bi-lateral rTMS vs Sham n=4,233 (number of patients in each group NR F/u: NR	OR: 3.03, 95% CI 1.19 to 7.71, p=0.02, favors bilateral rTMS	Yes (-1)	No	No	No	No	Moderate
Remission	81 RCTs included in 1 SR Brunoni et al., 2017 Note: does not indicate	Bi-lateral rTMS vs Sham n=4,233 (number of patients in each group NR	OR: 5.75; 95% CI, 1.93 to 17.24, favors rTMS	Yes (-1)	No	No	No	No	Moderate

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
	how many studies contributed to specific comparison	F/u: NR							
rTMS + Paroxetine vs Sham + Paroxetine for Acute Depression									
Depression symptoms	1 RCT Wang et al 2017	rTMS+ Paroxetine (22) vs Sham + Paroxetine (21) 4 weeks	rTMS MD: 8.41 ± 4.02; Sham MD: 11.29 ± 4.60, p<0.05, favors rTMS+ paroxetine	No	No	No	Yes (-2), very small sample size	No	Low
Response	1 RCT Wang et al 2017	rTMS (22) vs Sham (21) 4 weeks	95.5% rTMS vs. 71.4% Sham, p=0.041, favors rTMS + paroxetine	No	No	No	Yes (-1), very small sample size	No	Low
Remission	1 RCT Wang et al 2017	rTMS (22) vs Sham (21) 4 weeks	68.2% rTMS vs. 38.1% Sham, p=0.009, favors rTMS + paroxetine	No	No	No	Yes (-1), very small sample size	No	Low

CI: confidence interval; BSI: Beck Scale for Suicide Ideation; CAPs: Clinician-Administered PTSD Scale; f/u: follow-up; MD: mean difference; mos: months; NA: not applicable; NR: not reported; NS: not significant; OR: odds ratio; PTSD: post-traumatic stress disorder; RCT: randomized controlled trials; rTMS: repetitive TMS; SD: standard deviation; SMD: standardized mean difference; SR: systematic review; TMS: transcranial magnetic stimulation; wks: weeks; WMD: weighted mean difference

Table 2. GRADE Factors Used to Assess the Quality of a Body of Evidence

Evidence Category	Definition
Study Quality (Internal Validity or Risk of Bias)	Study quality considers the overall risk of bias rating of all the studies included in the evidence base. In this review, the overall risk of bias would be the average or median USPSTF rating for studies comprising an evidence base for a key outcome.
Consistency of Evidence	Consistency of evidence refers to the degree of similarity in the direction of effects or the degree of similarity in the effect sizes (magnitude of effect) across individual studies within an evidence base.
Directness of Evidence	Direct evidence directly compares interventions of interest in populations of interest and measures patient-oriented outcomes. Evidence can be indirect if the tested intervention differs from the intervention of interest, the study population differs from the population of interest, the outcomes differ from those of primary interest, or treatment comparisons have not been tested in head-to-head comparisons.
Precision of Evidence	Precision is the degree of certainty surrounding an estimate of effect with respect to an outcome. Precision is primarily assessed by examining the 95% confidence intervals around the summary effect size.

Link to GRADE Handbook: **http://gdt.guidelinedevelopment.org/app/handbook**

Table 3. Evidence Table for Systematic Reviews on Transcranial Magnetic Stimulation (TMS) to Treat MDD

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
Reference: Brunoni et al. 2017 *Weissman et al (2018) reports on suicidality of 2 studies included in this review that assess this outcome (Blumberg et al. 2012 & Blumberg et al. 2016) **Organization/Country:** Brazil **Purpose:** To establish the relative efficacy and acceptability of the different modalities of rTMS used for MDD by performing a network meta-analysis, obtaining a clinically meaningful treatment hierarchy **AMSTAR Rating:** Moderate **Overall RoB of Included Studies:** According to the authors (based on the Cochrane tool) 21.0%, 67.9%, and 11.1% of studies had an overall low, unclear, and high risk of bias; unclear rating due to lack of reporting of randomization or allocation procedures and/or imperfect blinding	**Databases Searched:** Searched PubMed/MEDLINE, EMBASE, PsycInfo, and Web of Science **Dates Searched:** Inception up until October 1, 2016 **Inclusion/Exclusion Criteria:** RCTs enrolling patients with a primary diagnosis of an acute unipolar or bipolar depressive episode, including studies of pts with comorbidities, such as anxiety or personality disorders. Studies must have compared at least 2 of the following interventions: LF-rTMS over the right DLPFC, HF-rTMS over the left DLPFC, bilateral rTMS (LF over the right and HF over the left DLPFC), TBS (including intermittent TBS over the left DLPFC, continuous over the right DLPFC or bilateral), pTMS over the right DLPFC, aTMS over the left DLPFC, sTMS, dTMS over the left DLPFC, and sham. Note: 1 Hz or less and 5 Hz or more defined low-frequency and high-frequency, respectively **Excluded** studies that enrolled participants with secondary mood disorders (e.g., post stroke depression); non-RCTs; trials performing less than 10 rTMS sessions; using frequencies between 2 to 4 Hz; or comparing	**Diagnosis:** MDD; most trials (74.1%) recruited only TRD patients **Number of Patients:** 4,233 **Age (mean years):** 46 **Gender (% female):** 59.1%	**Intervention:** LF-rTMS over the right DLPFC, HF-rTMS over the left DLPFC, bilateral rTMS (LF over the right and HF over the left DLPFC), theta-burst stimulation (TBS, including intermittent TBS over the left DLPFC, continuous over the right DLPFC or bilateral), pulsed-TMS over the right DLPFC, aTMS over the left DLPFC, synchronized TMS, deep TMS over the left DLPFC Monotherapy (10 RCTs); Add-on or augmentative therapy (71 RCTs) **Comparators:** Sham (as the most common comparator) **Follow-up:** NR **Outcomes: Primary**: Response (defined as 50% or greater improvement from baseline according to the study primary depression scale) and acceptability (measured as dropout) **Secondary:** Remission (defined as 7 or less, 8 or less, or 10 or less on the HDRS-17, HDRS-21, or MADRS, respectively)	**Response** **Direct evidence** **Bilateral rTMS vs. Sham:** OR, 3.39, 95% CI, 1.91 to 6.02, favors bilateral rTMS **HF-rTMS vs Sham**: OR, 3.28, 95% CI, 2.33 to 4.61, favor HF rTMS **LF-rTMS vs. Sham**: OR, 2.48, 95% CI, 1.22 to 5.05, favors LF-rTMS **TBS vs Sham**: OR, 2.57, 95% CI, 1.17 to 5.62, favors rTMS **Indirect evidence (NMA):** bilateral rTMS was more effective than sTMS (OR, 3.65; 95% CI, 1.02- 13.06) but no other important difference was found between the 8 active rTMS interventions. **Resolution of suicidal ideation** (defined as a decrease from any non-zero score at baseline to a zero score at endpoint on the HDRS-17 suicide item) **Bilateral rTMS vs. Sham:** OR, 3.03, 95% CI 1.19 to 7.71, p=0.02, favors bilateral rTMS

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
	only 1 modality of rTMS intervention **Final Evidence Base:** 81 RCTs			**Unilateral (left) rTMS vs. Sham:** OR, 1.59, 95% CI 0.61 to 4.12, p=0.33, NS **Acceptability** **Direct evidence:** Bi-lateral rTMS is more acceptable than LF-rTMS: OR, 2.43; 95% CI, 1.11-5.30) **Indirect evidence (NMA):** pTMS is significantly more acceptable than HF-rTMS (OR, 3.45; 95% CI, 1.15 to 10.0); LFrTMS (OR, 3.70; 95%CI, 1.25 to 11.1); sTMS (OR, 4.35; 95% CI, 1.3 to 14.29); and sham (OR, 3.70; 95% CI, 1.25 to 11.11) **Remission** **Direct evidence:** bi-lateral rTMS more effective than HF-rTMS (OR, 4.02; 95% CI, 1.3 to 12.35) and both interventions perform better than sham (OR, 5.75; 95% CI, 1.93 to 17.24 and OR, 2.72; 95% CI, 1.92 to 3.86, respectively) **Indirect evidence (NMA):** bilateral rTMS performs better than sTMS in terms of

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
				remission (OR, 4.95; 95% CI, 1.03 to 23.71) while bilateral (OR, 4.22; 95% CI, 1.96 to 9.05), LF-rTMS (OR, 2.70; 95% CI, 1.51 to 4.82), HF-rTMS (OR, 2.73; 95% CI, 1.78 to 4.20), and pTMS (OR, 4.37; 95% CI, 1.10 to 17.47) are more effective than sham **Limitations:** Review authors did not report on adverse events.
Reference: Health Quality Ontario, 2016 **Organization/Country:** Health Quality Ontario/ Canada **Purpose:** To examine the antidepressant efficacy of rTMS in patients with treatment-resistant unipolar depression. **AMSTAR Rating:** High **Overall RoB of Included Studies:** Some concerns; only 3 studies performed allocation concealment and not all studies blinded patients.	**Databases Searched:** Searched MEDLINE, MEDLINE In-Process and Other Non-Indexed Citations, Embase, EBSCO Cumulative Ontario Health Technology Assessment Series; Index to Nursing & Allied Health Literature (CINAHL), PsychInfo, and EBM. **Dates Searched:** January 1, 1994, to November 20, 2014. The search was updated on March 1, 2015, through the AutoAlert function of the search **Inclusion/Exclusion Criteria:** Included RCTs; Studies comparing rTMS with ECT or sham treatment in adult patients (age ≥ 18 years); Studies in which at least 80% of patients were resistant to treatment; Studies that applied high-frequency rTMS (≥5 Hz) to the	**Diagnosis:** Treatment resistant unipolar depression Note: While in most studies (n=16), patients had failed to benefit from two or more antidepressant medications, seven studies also included patients who had failed to improve with at least one antidepressant medication. **Number of Patients:** 1,156; n=602 rTMS; n=554 sham treatment. **Age:** Mean age range 39 to 64 yrs	**Intervention:** The frequency of stimulation in these studies ranged from 5 to 20 Hz, and the intensity of stimulation was between 80% and 120% of the patients' motor threshold. The number of trains per session ranged from 15 to 75, and train duration ranged from 2 to 10 seconds rTMS treatment was delivered over the course of 10 to 30 sessions. **Comparators:** Sham Note: In 16 studies, patients received rTMS while receiving antidepressants, and in seven studies patients did not receive any antidepressant during rTMS treatment. **Follow-up:** 2 wks to 6 mos **Outcomes: Primary**: Changes in depression scores measured by the HAM-D	**HF (≥5 Hz) Dorsolateral rTMS vs. Sham** **Depression Scores (17 studies):** WMD: 2.31 points, 95% CI 1.19–3.43, $p< 0.001$, favors rTMS, I^2=19.8%. The mean difference was below the mean value deemed a priori to be clinically important (i.e., the value of at least 3.5 points on the HAM-D). Meta-regression results indicated that frequency of stimulation, intensity of stimulation, and train duration were significantly associated with the treatment effect (p=0.002, 0.008, and 0.001, respectively)

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
	left dorsolateral prefrontal cortex and complied with rTMS safety guidelines; Studies that included unipolar patients only or that reported the proportion of bipolar patients as ≤20%; Studies in which patients received at least 10 sessions of rTMS treatment. **Excluded:** Non-RCTs; Studies of stimulation sites other than left dorsolateral prefrontal cortex; Studies that used frequencies of rTMS outside the range for this review; Studies on bilateral rTMS or on bilateral versus unilateral rTMS; Studies on sequential combined low-frequency and high-frequency rTMS; Studies on newer techniques (synchronized rTMS, pulsed rTMS, deep rTMS, rTMS with priming stimulation); Studies that evaluated the effect of rTMS on cognitive functions; Studies that evaluated the effectiveness of rTMS in depression due to specific conditions (i.e., poststroke depression, postpartum depression); Studies that did not report the important outcomes for this review, did not define the reported outcomes, or provided insufficient data **Final Evidence Base:** 23 RCTs comparing rTMS to sham	**Gender (% female):** 59.1%	**Secondary:** Remission rate measured by the HAM-D; Response rate measured by the HAM-D; Relapse rate; AEs	**Remission Rate (13 studies):** RR: 2.20, 95% CI 1.44–3.38, p< 0.001, I^2=0.0% **Response Rate (20 studies):** 1.72, 95% CI 1.13–2.62, p=0.011, I^2=46.4% There was a 10% difference in the rates of remission or response. This translates to a number needed to treat of 10. If 10 patients are treated with rTMS, one will have a chance to have a response or remission. **3 to 4 mos f/u (3 studies):** no evidence of difference between rTMS and sham **AEs:** Headache and scalp discomfort were the most frequently reported adverse events in these trials, and rates were higher in rTMS-treated than sham rTMS–treated patients No evidence of publication bias **Limitations:** Few studies (n=3) reporting on long-term effects of rTMS, which limits clear

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
				understanding of longer-term adverse events.

AE: adverse events; ATHF, Antidepressant Treatment History Form; BSI: Beck Scale for Suicide Ideation; CAPs: Clinician-Administered PTSD Scale; CI: confidence interval; CSSRS, Columbia Suicide Severity Rating Scale; DLPFC: dorsolateral prefrontal cortex; f/u: follow-up; HAM-A: Hamilton Anxiety scale; HAM-D: Hamilton Depression scale; HRSD, Hamilton Rating Scale for Depression; HF-rTMS: high frequency rTMSI^2: % of heterogeneity between studies; LF-rTMS: low frequency rTMS; mos: months; MADRS: Montgomery-Asberg Depression Rating Scale; MDD: major depressive disorder; NMA: network meta-analysis; NR: not reported; NS: not significant; OR: odds ratio; PTSD: post-traumatic stress disorder; RCT: randomized controlled trials; RoB: risk of bias; RR: risk ratio; SCID, Structured Clinical Interview for DSM Disorders; SMD: standardized mean difference; SR: systematic review; rTMS: repetitive TMS; TMS: transcranial magnetic stimulation; TRD: treatment resistant depression; wks: weeks; WMD: weighted mean difference

Table 4. Systematic Review Risk of Bias AMSTAR Checklist Table on TMS to Treat MDD

Question	Brunoni et al. 2017	Health Quality Ontario, 2016
Did the research questions and inclusion criteria for the review include the components of PICO?	Yes	Yes
Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	Yes	Yes
Did the review authors explain their selection of the study designs for inclusion in the review?	Yes	Yes
Did the review authors use a comprehensive literature search strategy?	Yes	Yes
Did the review authors perform study selection in duplicate?	Yes	Yes
Did the review authors perform data extraction in duplicate?	Yes	Yes
Did the review authors provide a list of excluded studies and justify the exclusions?	No	Yes
Did the review authors describe the included studies in adequate detail?	Yes	Yes
Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	Yes	Yes
Did the review authors report on the sources of funding for the studies included in the review?	No	Yes
If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results? RCTs?	Yes	Yes
If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	Yes	Yes
Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	Yes	Yes
Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	Yes	Yes
If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	Yes	Yes
Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	Yes	Yes
Overall Quality	Moderate	High

RoB: risk of bias

Table 5. AMSTAR Rating of Overall Confidence in Results of the Review

Category	Definition
High	*No or one non-critical weakness*: the systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest.
Moderate	*More than one non-critical weakness*: the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review.
Low or Very Low	*One or more critical flaw(s) with or without non-critical weaknesses*: the systematic review has one or more critical flaws and may not provide an accurate and comprehensive summary of the available studies that address the question of interest.

AMSTAR checklist, go to https://amstar.ca/Amstar_Checklist.php

Table 6. Evidence Table for RCTs on Transcranial Magnetic Stimulation (TMS) to Treat MDD

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
Reference: Yesavage et al 2018 **Purpose:** To determine the efficacy of rTMS in the treatment of TRD in veterans **Setting:** 9 VA medical centers; primary Palo Alto **Funding source:** VA	**Number of patients:** 164; n=81 rTMS; n=83 sham **Inclusion criteria:** Between ages 18 and 80 yrs with a DSM-IV diagnosis of MDD; HRSD score 20 no more than 7 days prior to randomization; Exhibit moderate level of resistance to antidepressant treatment defined, using the ATHF, as failure of 2 adequate medication trials; Duration of current episode of MDD 10 yrs; Ability to obtain a motor threshold (should be determined at the end of the screening process); Currently under the care of a VA psychiatrist; If receiving a psychotropic medication regimen, that regimen will be stable for 4 weeks prior to randomization and patient will be willing to continue receiving a stable regimen during the acute treatment phase; Women agree to use acceptable methods of birth control; and ability to read and understand consent form. **Exclusion criteria:** Pregnancy or lactation; Unable to be safely withdrawn from medications that increase the risk of having seizures; Cardiac pacemaker; Implanted device (deep brain stimulation) or metal in the brain; Cochlear implant; Mass lesion, cerebral infarct, increased intracranial pressure, or other active central nervous system disease, including a seizure disorder; Known current psychosis, bipolar, amnestic disorders,	**Intervention:** Left prefrontal rTMS treatment (10 Hz, 120% motor threshold, 4000 pulses/ session) for up to 20 to 30 sessions **Control:** Sham (control) rTMS treatment for up to 30 treatment sessions. For both groups, treatment was delivered in 5 session blocks over a period of 5 to 12 calendar days. **Outcomes:** Remission rate (HRSD score 10, indicating that depression is in remission and not a clinically significant burden); indices of PTSD, depression, hopelessness, suicidality, and quality of life; and AEs **F/u:** posttreatment at 5 to 12 days and 24 weeks	**Remission of depression** (OR, 95% CI, p-value): **Posttx**: 1.16, 95% CI 0.59 to 2.26, p=0.67 **Achieved remission** rTMS: 33/81 (40.7%) vs. sham 31/83 (37.4%) **24 wks:** 1.55, 95% CI 0.62 to 3.86, p=0.35 **Sustained remission** rTMS: 16 (19.8%) vs sham 13 (15.7%) **Depression symptom severity** (HRSD. MD, 95% CI, p-value) **Posttx:** 1.28, 95% CI -1.42 to 3.97, p=0.34 **24 wks:** 0.62, 95% CI -2.59 to 3.94, p=0.68 **Suicidality** (BSI, MD, 95% CI, p-value) **Posttx:** 0.08, 95% CI -1.46 to 1.62, p=0.91 **24 wks:** -0.54, 95% CI -2.25 to 1.17, p=0.53 **PTSD symptom severity** (CAPS, MD, 95% CI, p-value) **Posttx:** 5.20, 95% CI -0.49 to 10.89, p=0.07	Results suggest that there was no significant difference between rTMS and sham in the number of patients in the acute phase of depression who experienced remission. Overall, 39% of the total study population experienced remission. There was also no difference between groups in severity of depression or PTSD symptoms, suicidality or quality of life. The most common AE in both groups was headache. **Limitations:** Limited reporting on concurrent treatment, such as number of pts receiving psychotropic medication or therapy; relatively short follow-up to measure long-term effects or AEs. **Study ROB:** Low **Author conflict:** None reported

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
	dementia, or other cognitive disorders; Current substance abuse (not including caffeine or nicotine); Elevated risk of seizure due to TBI; Prior exposure to rTMS; or Active current suicidal intent or plan **Pt. baseline characteristics (all pts):** Age (mean, yrs.): 55.2 yrs Gender (% male): 80.5% HRSD: 26.9 (5.0)		**24 wks:** 4.47, 95% CI -0.69 to 9.64, p=0.09 **Quality of life** (Veterans RAND-36, MD, 95% CI, p-value) **Posttx** (Physical): -1.32, 95% CI -3.61 to 0.97, p=0.27 (Mental): -1.76, 95% CI -5.91 to 2.39), p=0.40 **24 wks** (Physical): 0.08, 95% CI -2.67 to 2.83, p=0.96 (Mental): -0.12, 95% CI -4.48 to 4.24, p=0.96 **Common non-serious AEs** **Nasopharyngitis:** 8 rTMS; 8 sham **Depression:** 8 rTMS; sham **Falls:** 3 rTMS; 7 sham **Headache:** 15 rTMS; 16 sham **Abnormal hearing tests:** 18 rTMS; 18 sham (believed to be an artifact of frequent, imprecise testing). **Common serious adverse** **Suicidal ideation:** 3 rTMS; 4 sham No suicides or seizures occurred during the	

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
			study and there were no deaths.	
Reference: Wang et al 2017 **Purpose:** To examine the effectiveness of rTMS on first episode depressed patients when combined with antidepressant drugs. **Setting:** **Funding source:** None stated	**Number of patients:** 43; n=22 rTMS; n=21 Sham **Inclusion criteria:** Patients ≥18 years who met DSM-IV diagnostic criteria for first episode of MDD. **Exclusion criteria:** (i) age < 18 years or > 45 years, (ii) comorbid DSM-IV diagnosis, including alcohol or illicit drug abuse, and other Axis I psychiatric disorders, (iii) current or past serious physical illness (e.g., active tuberculosis, acute hepatitis, cirrhosis, renal illness, cardiovascular illness, or unstable diabetes), (iv) diastolic blood pressure < 60 mm Hg or heart rate < 60 beats per minute, (v) HDRS suicidal ideation score ≥3, (vi) risk factors for the rTMS procedure (e.g., epilepsy, severe and repetitive headache episodes, previous neurosurgery, implants of metal or clips, and pregnancy), and (vii) those female participants who were pregnant or were planning to become pregnant. **Pt. baseline characteristics (rTMS; Sham):** Age (mean, yrs.): 28.82 ± 8.46; 30.05 ± 9.47 Gender (% male): 54.5%; 47.6% Mean duration of current episode (months) 4.32 ± 2.25; 4.31 ± 1.83 HRSD score Baseline 43.50 ± 9.89; 42.81 ± 9.29	**Intervention:** The stimulation parameters in the rTMS group were the following: 10 Hz, 80% of motor threshold over the left dlPFC, and 40 trains of 20 pulses (2 s each, with a 28-s intertrain interval). A total of 800 pulses were administered per day, with 20 sessions (five sessions per week) for 4 weeks. Paroxetine dosage began with 10 or 20 mg daily (taken after breakfast) for the first week and then was titrated to 20 or 30 mg daily from day 8 onward. **Control:** Sham + paroxetine **Outcomes:** Severity of depressive symptoms using the HDRS-24; Remission; and AEs **Note:** Pts with a 50% reduction in the absolute HDRS-24 score at the end of the 4th and 8th week of treatment from baseline were classified as responders and were considered remitted with a HDRS-24 residual score < 8 **F/u:** 8 weeks	**4 weeks** **Depression symptoms (HDRS scores):** rTMS 8.41 ± 4.02; Sham 11.29 ± 4.60, p<0.05, favors rTMS+ paroxetine **Week 8** rTMS 7.32 ± 3.24; Sham 8.14 ± 4.50 p=0.48, NS **4 weeks** **Response rate**: 95.5% rTMS vs. 71.4% Sham, p=0.041, favors rTMS + paroxetine **Remission rate:** 68.2% rTMS vs. 38.1% Sham, p=0.009, favors rTMS + paroxetine **8 weeks** **Response rates:** 90.9% rTMS vs. 85.7% Sham, p=0.189, NS **Remission:** 86.4% rTMS vs. 76.2% Sham, p=0.069, NS **AE's** Commonly reported AEs: Headache or scalp pain: 5 rTMS; 7 sham No seizures, hearing impairment, or	Results suggest that rTMS plus paroxetine improved symptoms of depression and led to higher response and remission rate compared to sham rTMS plus paroxetine after 4 weeks of treatment. However, no differences were observed between the active and sham rTMS at 8 weeks follow-up. Headache and scalp pain were the most commonly reported adverse event with no reported incidence of seizures, hearing loss or memory problems. **Limitations:** Small sample size and limited follow-up **Study ROB:** Low **Author conflict:** None stated

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
			subjective complaints about memory or concentration reported among pts in rTMS.	

AE: adverse events; ATHF, Antidepressant Treatment History Form; BSI: Beck Scale for Suicide Ideation; CAPs: Clinician-Administered PTSD Scale; CI: confidence interval; CSSRS, Columbia Suicide Severity Rating Scale; DLPFC: dorsolateral prefrontal cortex; f/u: follow-up; HAM-A: Hamilton Anxiety scale; HAM-D: Hamilton Depression scale; HRSD, Hamilton Rating Scale for Depression; HF-rTMS: high frequency rTMSI^2: % of heterogeneity between studies; LF-rTMS: low frequency rTMS; mos: months; MADRS: Montgomery-Asberg Depression Rating Scale; MDD: major depressive disorder; NMA: network meta-analysis; NR: not reported; NS: not significant; OR: odds ratio; PTSD: post-traumatic stress disorder; RCT: randomized controlled trials; RoB: risk of bias; RR: risk ratio; SCID, Structured Clinical Interview for DSM Disorders; SMD: standardized mean difference; SR: systematic review; rTMS: repetitive TMS; TMS: transcranial magnetic stimulation; TRD: treatment resistant depression; wks: weeks; WMD: weighted mean difference

Table 7. Cochrane Risk of Bias 2.0 Tool for RCTs on TMS to Treat MDD

Reference	Yesavage et al. 2018	Wang et al. 2017
• Was the allocation sequence generated adequately (e.g., random number table, computer-generated randomization)?	Yes	Yes
• Was the allocation of treatment adequately concealed (e.g., pharmacy-controlled randomization, concealed envelopes)?	Yes	Yes
• Did baseline difference between study groups suggest a problem with randomization?	No	No
Overall RoB for Randomization Process	**Low**	**Low**
Deviation from Intended Intervention (Effect of Assignment)		
• Were participants aware of their assigned intervention during the trial?	No	No
• Were providers and people delivering treatment aware of assigned intervention during trial?	No	Yes
• Were there deviations from the intended intervention that arose because of the experimental context?	No	No
• Were these deviations from intended intervention balanced between groups?	No	No
• Were these deviations likely to have affected the outcome?	No	No
• Was an appropriate analysis used to estimate the effect of assignment to intervention?	No	No
Overall RoB of Effect of Assignment	**Low**	**Low**
Missing Outcome Data		
• Were data for this outcome available for all, or nearly all, participants randomized?	Yes	Yes
• Is there evidence that result was not biased by missing outcome data?	NA	NA
• Could missingness in the outcome depend on its true value?	NA	NA
• Do the proportions of missing outcome data differ between intervention groups?	NA	NA
• Is it likely that missingness in the outcome depended on its true value?	NA	NA
Overall RoB of Missing Data	**Low**	**Low**
Measurement of the Outcome		
• Was the method of measuring the outcome inappropriate?	No	No
• Could measurement or ascertainment of the outcome have differed between intervention groups?	No	No
• Were outcome assessors aware of the intervention received by study participants?	No	No
• Could assessment of the outcome have been influenced by knowledge of intervention received?	No	No
• Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	No	No
Overall RoB of Measurement of Outcome	**Low**	**Low**

Reference	Yesavage et al. 2018	Wang et al. 2017
Selection of Reported Results		
• Was the trial analyzed in accordance with a pre-specified plan that was finalized before unblinded outcome data were available for analysis?	Yes	Yes
Overall RoB of Reported Results	**Low**	**Low**
Overall Study RoB	**Low**	**Low**

*Responses: Y=Yes; PY=Probably Yes; N=No; PN=Probably No; NA=Not Applicable; NI=No Information; RoB: risk of bias

Table 8. Cochrane Risk of Bias 2.0 Overall Risk of Bias Judgement

Category	Definition
Low risk of bias	The study is judged to be at low risk of bias for all domains for this result.
Some concerns	The study is judged to be at some concerns in at least one domain for this result.
High risk of bias	The study is judged to be at high risk of bias in at least one domain for this result. OR The study is judged to have some concerns for multiple domains in a way that substantially lowers confidence in the result.

References

Brunoni, A. R., Chaimani, A., Moffa, A. H., Razza, L. B., Gattaz, W. F., Daskalakis, Z. J., & Carvalho, A. F. (2017). Repetitive transcranial magnetic stimulation for the acute treatment of major depressive episodes a systematic review with network meta-analysis. *JAMA Psychiatry, 74*(2), 143–152. https://doi.org/http://dx.doi.org/10.1001/jamapsychiatry.2016.3644

Health Quality Ontario (2016). Repetitive transcranial magnetic stimulation for treatment-resistant depression: a systematic review and meta-analysis of randomized controlled trials. *Ontario Health Technology Assessment Series, 16*(5), 1-66.

Wang, H. N., Wang, X. X., Zhang, R. G., Wang, Y., Cai, M., Zhang, Y. H., … Zhang, Z. J. (2017). Clustered repetitive transcranial magnetic stimulation for the prevention of depressive relapse/recurrence: A randomized controlled trial. *Translational Psychiatry, 7*(12), 1292. https://doi.org/https://dx.doi.org/10.1038/s41398-017-0001-x

Yesavage, J. A., Fairchild, J., Mi, Z., Biswas, K., Davis-Karim, A., Phibbs, C. S., … George, M. S. (2018). Effect of repetitive transcranial magnetic stimulation on treatment-resistant major depression in US veterans: A randomized clinical trial. *JAMA Psychiatry, 75*(9), 884–893. https://doi.org/http://dx.doi.org/10.1001/jamapsychiatry.2018.1483

Exercise

Evidence Base

Our searches of the literature identified 1 SR and 9 RCTs that assessed the impact of integrating exercise interventions into care for the treatment of adults with MDD (See **Tables 3 and 6** for details on study characteristics). The SR by Kvam et al. (2016) was a meta-analysis of RCTs that examined the efficacy of physical exercise as a treatment of depression, both as an independent intervention and as an adjunctive treatment to pharmacotherapy for depression. Physical exercise could be aerobic (e.g. walking, running or cycling) or non-aerobic (e.g. resistance, weightlifting, or strength). Comorbid conditions were allowed if the primary mental health condition was unipolar depression. The main outcome measure needed to be depression symptoms measured by a validated scale. A total of 23 RCTs and 977 subjects were included in the study. Physical exercise had a large to moderate statistically significant effect on depression when compared to control, but the effect size was small and not significant at follow-up. Exercise compared to no intervention also had a large and significant effect. Exercise compared to TAU had a moderate and significant effect. However, the effects of exercise compared to either behavioral treatments or pharmacotherapy were small and not significant. Exercise as an adjunct to behavioral treatments had a moderate but nonsignificant effect. When analysis was restricted to the six studies with methods that included allocation concealment, blinded outcomes and intention to treat analysis, the effects of depressive symptoms were small and no longer significant.

The trial by Abdollahi et al. (2017) randomized 54 mildly to moderately depressed individuals (54% women; average age 48.25 years old) to receive either combined (adjunctive) CBT and exercise or CBT only. Both groups received one session of CBT weekly for 12 weeks while the CBT/exercise group also completed exercise 3 times each week supervised by a PhD in trained in sports science. Exercise consisted of warm up, cardiovascular exercise, walking and cool down at moderate intensity. Multilevel modeling demonstrated greater improvement in suicidal ideation, depression, and activities of daily living in the combined CBT/exercise group when compared to CBT only. The study was limited by drop-outs and lack of follow-up.

Belvederi-Murri et al. (2015) studied the effects of augmenting sertraline treatment with exercise in late-life depression. The investigators randomized 121 primary care patients > 65 years who had MDD to one of three groups: 1) higher-intensity, progressive aerobic exercise plus sertraline (S+PAE); 2) lower-intensity, non-progressive aerobic exercise plus sertraline (S+NPAE); or 3) sertraline only, for 24 weeks. At the end of the study, a significantly higher number of individuals in the exercise groups achieved remission; 45% of the sertraline group, 73% of the S+NPAE group, and 81% of the S+PAE group attained remission (as measured by a HSRD score of $\leq$10).

Another trial by Combs et al. (2014) studied exercise as a monotherapy for major depressive disorder as diagnosed by the SCID and meeting DSM VI criteria. This study randomized 202 adults to one of four arms: 1) supervised aerobic exercise, 2) home-based supervised exercise that identically matched the intensity and frequency of the supervised exercise group, 3) sertraline treatment, or 4) placebo pill. The aim of this study was to assess the effects of either exercise or sertraline on disordered sleep in adults with MDD. Participants were followed for 12 months. The study demonstrated that the active treatment and placebo groups showed no differences in the HAM-D sleep components after 4 months. Neither sertraline

nor exercise was associated with greater improvements in sleep when compared with placebo. However, residual symptoms of insomnia after successful MDD treatment predicted relapse.

In Doose et al. (2015), the purpose of the study was to determine the effect of physical exercise on unipolar depression. The investigators randomized 46 adult outpatients, aged 18 – 65 years old, who met ICD-10 criteria for mild to severe MDD. All but 3 had recurrent MDD; 58.7% received concurrent psychotherapy or pharmacotherapy. Participants were randomized to a walking/running aerobic exercise program at the local sports club or to a wait list control group. The study period was eight weeks and involved exercise three times in a one-hour group session. The study was limited by a 24% drop out and a 58% session attendance. There was a large and clinically significant decrease in depressive symptoms measured by HSRD- 17 but BDI –II and VO_2 did not change significantly. The intervention group, as measured by HSRD-17, was considered "recovered" while the control group was described as "unchanged".

Kerling et al. (2015) studied 42 inpatients in German hospitals with moderate to severe depression. The aim of the study was to examine whether exercise as an adjunct to inpatient treatment improved psychological and physiological factors. Twenty-two patients were randomized to exercise (cycling and crosstraining, stepping, upper body, treadmill, recumbent, or rowing) three times per week for six weeks compared to twenty individuals who received TAU (CBT and pharmacotherapy). The investigators reported that there were no dropouts and a 90% participation rate. This study found that exercise improved physical fitness (VO2 and ventilatory anaerobic threshold). Both groups showed improvements in depressive symptoms (BDI-II and MADRS), however significantly more patients in the exercise group were classified as responders. Changes in BDI-II, MADRS sum score, and remission were not significant.

In a study by LeGrand et al. (2016), the effects of physical exercise as adjunctive treatment in patients hospitalized for MDD (DSM-IV and score of > 29 on BDI – II). Patients on antidepressants for < 2 weeks were included in the study. Thirty-five patients (mean age 45.3 $\pm$ 13.2 yrs.; 71% women) were randomized to one of three groups 1) aerobic exercise (30 minutes of brisk walking or running); 2) placebo stretching exercises; or 3) no treatment. They found a large and significant effect on BDI - II in both the aerobic exercise and stretching group but no significant change in depressive symptoms in the controls.

Schuch et al. (2015) performed an RCT to evaluate the effects of adjunctive exercise in hospitalized patients with severe MDD. Fifty patients (were randomized to exercise plus TAU (n =25) or TAU control (n = 25). 52.8% of patients refused to participate in the trial. The twenty-five patients randomized to the exercise (stretching, walking on a treadmill, and exercise of patient's preference) group did three exercise sessions per week throughout the period of hospitalization. Depression symptoms and quality of life (QoL) were evaluated at baseline, 2 weeks, and discharge. Significant differences in depression symptoms and QoL (both physical and psychological) at 2 weeks and at discharge. However, there was no significant difference in the response or remission rates when the exercise group was compared to usual care.

Siquiera et al. performed a randomized single-blind study that aimed to evaluate the effect of aerobic exercise as an adjunct to antidepressant therapy. Fifty-seven patients (41 women with a mean age of 38.82$\pm$ 10.72 years) were randomized to either a 4 week, 4 times/week program of aerobic exercise or no activity. Depression severity measured using the HAM-D and BDI. The dropout rate was 29.8%.

Depression severity decreased significantly in both groups but there were no differences between the exercise and usual care groups. The authors noted that the exercise group required lower doses of sertraline monotherapy.

In the final RCT, Verrusio et al. randomized 24 elderly subjects with mild to moderate MDD according to DSM-IV criteria (mean age 75.5 $\pm$ 7.4, 11 M and 13 F) into a pharmacotherapy group and exercise combined with listening to music. The authors stated that all subjects completed the study. There was a significant reduction of depression and anxiety at 12 and 24 weeks compared to baseline in the exercise/music listening group ($p = 0.05$). While the pharmacotherapy group showed a reduction in anxiety only at 24 weeks.

Study Quality

We rated the strength of evidence of the individual RCTs as Low or Very Low due to concerns regarding risk of bias due to lack of information on allocation concealment, lack of blinding of patients, lack of blinding of some clinicians and/or outcome assessors, and attrition. In addition, small sample size and short study duration was characteristic of many studies. In most studies, long term outcomes were not available. (See **Table 7** for the RoB ratings). We rated the Kvam et al. meta-analysis as being of Moderate quality using AMSTAR criteria (see **Table 4** for the rating).

Key Findings

Below, we describe the key findings for the outcomes of interest with the GRADE strength of the evidence (SOE) rating. See **Table 1** for factors that influenced the SOE ratings.

- Evidence from 1 systematic review with 23 RCTs suggest that exercise statistically significantly reduces depressive symptoms immediately following treatment when compared to no treatment or WL control (SOE: Low)
- Evidence from 3 RCTs suggest that adjunctive aerobic exercise therapy of any kind (e.g., running, cycling, stretching, or walking) statistically significantly improves symptoms of depression in outpatients compared to controls (e.g., WL, TAU, or other active treatment with pharmacotherapy or psychotherapy). Results were inconsistent as 1 RCT showed no significant effects. (SOE: Low)
- Evidence from 3 RCTs suggest that adjunctive aerobic exercise therapy statistically significantly improves symptoms of depression in inpatients compared to controls (e.g., WL, TAU, or other active treatment with pharmacotherapy or psychotherapy). 1 RCT showed increased remission at discharge for aerobic exercise compared to controls. (SOE: Low)
- Evidence from 1 RCT suggests that aerobic exercise of any kind (e.g., aerobic, strength training, sailing) statistically significantly improves psychological quality of life compared to WL. (SOE: Low)
- Evidence from 1 RCT of mild to moderate depression suggests that the addition of exercise to CBT produces a short-term reduction of suicidal ideation compared to CBT alone. Long term follow-up studies were not available. (SOE: Very Low).

- Evidence from 1 RCT suggests that exercise alone may reduce symptoms of anxiety compared to controls. (SOE: Very low)
- Evidence from the secondary analysis of 1 RCT suggests that there is no statistically significant difference in sleep quality between exercise therapy plus treatment as usual and treatment as usual alone. (SOE: Very low)

Discussion

Overall, the findings of a single SR and the 9 RCTs that made up the evidence base for exercise suggest that exercise (e.g., aerobic, stretching, or walking) when used as an adjunct to medication and/or psychotherapy reduces symptoms of depression compared to controls (e.g., waitlist, TAU, or other active treatment). One study in elderly individuals with mild to moderate depression demonstrated that exercise and listening to music alone reduced both depression and anxiety when compared to medication alone. See **Table 1** for a summary of all the findings for exercise to treat adults with MDD. The strength of the evidence supporting the findings for exercise in reducing symptoms of depression, anxiety, and suicidal ideation was rated as low or very low mainly due to methodological limitations of the included studies. These limitations included lack of blinding of participants, clinicians and outcome assessors and high rates of attrition and nonadherence. However, limited evidence (1 RCT each) suggest that there is no significant difference between exercise and controls in improving psychological quality of life or sleep quality.

Table 1. Strength of Evidence for Exercise to Treat MDD

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
Any Type of Exercise vs. Any Control									
Depression Symptoms	1 SR w/ 23 RCTs in Kvam; 7 RCTs	Exercise; includes aerobic, strength training, stretching, cervical exercises, and sailing Control; includes No tx., WL, TAU, or standard PT	Kvam: Exercise vs. control: (g = -0.68 (95% CI = -0.92 to -0.44), p= 0.001); **Favors exercise** Exercise vs. control F/U: g=-0.22 (95%CI=-0.53 to 0.09, p=0.16; **NS** Exercise vs. No intervention: g=-0.48 (95%CI -0.80 to 0.16, p< 0.001); **Favors exercise** Exercise vs. Psychotherapy: g=-0.22 (95% CI = -0.65 to 0.21, p = 0.31; **NS** Exercise + Antidepressants vs. Antidepressants alone: g = -0.08 (95%CI = -1.10 to 0.11, p=0.11); **NS** Abdollahi: (CBT/group aerobic exercise vs. CBT only SD 2.25, r^2=0.84, 95%CI= [1.43, 2.83]	Yes (-2)	No	No	No	NA	Low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
			Favors adjunctive exercise Belvederi – Murri (Sertraline plus progressive aerobic [S+PAE] vs sertraline plus lower intensity non-progressive aerobic exercise [S+NPE]) HSRD 4 wks.: Both exercise groups showed improvement over sertraline alone 2LL decreases > 45, d.f. =3, $p<0.01$. Pt rated "much improved": S-PAE – 71%, S+NPE – 60%, S – 43%, $X^2 = 7.09$. d.f. = 2, p=0.03 **Favors adjunctive exercise** Doose et al. 2015 Aerobic exercise as an adjunct to TAU (n = 30) vs. waitlist (n=16) HSRD - 8.24 (p=,0.0001, 95%CI [-11.45, -5.02]						

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
			BDI –II: -8.20 (95%CI: -11.39; -5.010 **Favors adjunctive exercise**						
			Kerling 2015 N=42 inpatients Exercise/CBT vs. TAU/CBT Both groups showed a decline in depressive symptoms but the exercise vs. TAU groups were not significantly different. MADRS – F=2.23; p=0.14 BDI-II: F=.69; p=0.41 **Not Significant**						
			Legrand 2015 Exercise as an adjunct to pharmacotherapy. (n=35; aerobic – 14; stretching – 11; no intervention – 10). Both the aerobic and (MSD=18.92 vs. 36.14, p=0.001) and stretching (MSD=28.43vs. 37.82, p=0.011)						

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
			resulted in significant improvement. While no intervention resulted in no change. Reduction was larger in the aerobic vs. NI group, MSD=-17.22vs. -6.41, p=0.012, and marginally larger for stretching-17.22 vs. 9.39, p=0.082. Stretching vs. NI was NS. The effect size for aerobic vs. control was large (Cohen's d = -1.39). The effect size of stretching vs. NI was near 0. (Cohen's d = -0.33) **Favors adjunctive exercise**						
			Schuch et al. 2015. Adjunctive aerobic exercise in inpatients with MDD. Exercise (n=25) vs. TAU (pharmacotherapy +/- ECT(n=25).						

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
			HAM-D 2 wk.: 4.41(95%CI -7.57-1.25; p=0.007) At discharge: 3.70 (95%CI 6.21-1.19; p =0.005) **Favors adjunctive exercise**						
			Siqueira et al. Exercise/SSRI vs. SSRI n=57 4 week: HAM-D no significant difference. BDI-II no significant difference **Not Significant**						
Quality of Life (social/ psychological)	1 RCT Schuch et al 2015	Aerobic Exercise vs. TAU	Psychological QoL, 2 wk. 12.99(95%CI 1.68-24.29, p = 0.029 At discharge: 19.10(95%C I9.58-28.62, p=0.01	Yes (-2)	No	No	No	NA	Low
			Social QoL: 2 wk: 67.68 At discharge: 63.62						
Quality of Life (physical)	1 RCT Schuch et al 2015	Aerobic Exercise vs. TAU	Physical QoL 2wk 15.64 (95%CI 6.05-25.23, p = 0.002) At discharge: 19.10 (95%CI9.58-28.62, p=0.01)	Yes (-2)	No	No	Yes (-1); wide 95% CIs	NA	Low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
MDD Remission	2 RCTs; Schuch 2015; Murri, 2015	Exercise; includes sertraline plus progressive aerobic exercise (S+PAE) and Sertraline plus low intensity/ non-progressive (S+NPE) vs. sertraline alone (S)	Schuch et al. 2015 48% exercise vs. 32% TAU; p=0.248. Number needed to treat = 6.25 Murri 2015 S+PAE= 42, N+NPE=37, S=42 **Remission rate:** 4 wk. = 36%, 40%, 7%, respectively (p=0.001) 8 wk. = 60%, 49%, 40% 9 (p=.22) 12 weeks: 83%, 54%, 45% (p=0.001 24 wk. =81%, 73%, 45% (p=0.001) **Time to Remission:** X2=12.6, d.f =2, p=0.002; shorter time to remission for S+PAE (9.3 wk, 95% CI 7.4 – 11.2) 12 wk.: -3.92[95%CI -7.50-0.33] (p=0.03) 24 wk.: -5.50 [-8.80-2.20] (p=0.00)	Yes (-2)	No	No	No	NA	Low
Sleep disturbance	1 RCT Combs, 2014	supervised exercise (SE), home-based	16 wk. All treatments showed improvement	Yes (-2)	No	No		NA	Low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
(HAM-D, 3 items)		exercise (HE), sertraline (S), placebo (P) (n=202)	(p=0.004); Active treatments were no better than placebo (p=0.867) NS Exercise vs sertraline were comparable (p=0.841) Residual insomnia symptoms were associate with increased depression at 1 yr. (beta = 0.26, p=0.005)						
Exercise plus Music Listening vs. Pharmacotherapy in Elderly with Mild to Moderate Depression									
Depression Symptoms	1 RCT Verrusio et al. 2014		Exercise/music therapy vs. pharmacotherapy N=24; age = 75.5+7.4 Depression measured by GDS 12 wk. -1.2[-2.84 to 0.84] (p=0.3) 24 wk. -2.92[-4.39 to -0.61] (p=0.01) **Favors exercise at 24 weeks**	Yes (-2)	No	No	Yes	NA	Very Low
Anxiety	1 RCT Verrusio et al. 2014	Exercise/music therapy vs. pharmacotherapy	Anxiety measured by HAS 12 wk.: -3.92[95%CI -7.50-0.33] (p=0.03)	Yes (-2)	No	No	Yes (-1); not significant	NA	Very low

Outcome	Quantity and Type of Evidence	Intervention (n)/ Control (n)/Follow-up	Estimate of Effect	Study Limitations (Risk of Bias)	Inconsistency	Indirectness	Imprecision	Publication Bias	GRADE of Evidence for Outcome
			24 wk.: -5.50 [-8.80- 2.20] (p=0.00) **Favors exercise at 12 & 24 weeks**						

BDI-II: Beck Depression Inventory-II; CBT: Cognitive Behavior Therapy; CG: control group; CI: confidence interval; ES: effect size; EX: exercise; f/u: follow-up; HAM_D: Hamilton Rating Scale for Depression; HAS: Hamilton Anxiety Scale; Geriatric Depression Scale: GDS; MDD: Major Depressive Disorder; NR: not reported; NS: not significant; RCT: randomized controlled trials; RoB: risk of bias; SD: standard deviation; TAU: treatment as usual; WL: waitlist control

Table 2. GRADE Factors Used to Assess the Quality of a Body of Evidence

Evidence Category	Definition
Study Quality (Internal Validity or Risk of Bias)	Study quality considers the overall risk of bias rating of all the studies included in the evidence base. In this review, the overall risk of bias would be the average or median USPSTF rating for studies comprising an evidence base for a key outcome.
Consistency of Evidence	Consistency of evidence refers to the degree of similarity in the direction of effects or the degree of similarity in the effect sizes (magnitude of effect) across individual studies within an evidence base.
Directness of Evidence	Direct evidence directly compares interventions of interest in populations of interest and measures patient-oriented outcomes. Evidence can be indirect if the tested intervention differs from the intervention of interest, the study population differs from the population of interest, the outcomes differ from those of primary interest, or treatment comparisons have not been tested in head-to-head comparisons.
Precision of Evidence	Precision is the degree of certainty surrounding an estimate of effect with respect to an outcome. Precision is primarily assessed by examining the 95% confidence intervals around the summary effect size.

Link to GRADE Handbook: **http://gdt.guidelinedevelopment.org/app/handbook**

Table 3. Evidence Table for Systematic Reviews on Exercise to Treat MDD

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
Reference: Kvam 2016 **Organization/Country:** University of Bergen, Norway **Purpose:** Meta-analysis to examine the efficacy of physical exercise as treatment for unipolar depression, both as an independent intervention and as an adjunct intervention to antidepressant medication **AMSTAR Rating:** Moderate **Overall RoB of Included Studies:** High	**Databases Searched:** PsycINFO, EMBASE, MEDLINE, CENTRAL, and Sports Discus **Dates Searched:** through 2014 **Inclusion/Exclusion Criteria:** **Inclusion:** RCTs of diagnosis of unipolar depression (DSM or ICD) in any setting. **Exclusion:** Non- randomized studies, seasonal depression, Bipolar disorder, children **Final Evidence Base:** 23 studies	**Diagnosis:** Unipolar depression (DSM-IV or ICD; minor depression or dysthymia) **Number of Patients:** 977 **Age:** 18 years or older **Gender:** Male and female	**Intervention:** Aerobic Exercise, Resistive Exercise **Comparators:** no treatment, pharmacotherapy (as monotherapy), psychotherapy (as monotherapy), Adjunct to pharmacotherapy or psychotherapy compared to **Follow-up:** variable but as short as 4 weeks **Outcomes:** Primary outcome was a validated score of depression symptoms or remission.	**Exercise vs. Control** 23 studies of 977 participants, found that reduction in depressive symptoms after treatment showed a moderate to large and significant effect in favor of exercise (g = -0.68 (95% CI = -0.92 to -0.44), p= 0.001).The heterogeneity between studies was significant and moderate to high(Q (22)=68.737, p<0.001, I^2=67.99). **Favors exercise** **Exercise vs. Control: F/U** In 7 studies of 348 participants, there was a small but NS effect on depressive symptoms, g=-0.22 (95%CI=-0.53 to 0.09, p=0.16). Heterogeneity between studies was moderate but NS. **Exercise vs. No Intervention** Four studies of 77 participants showed a large and significant effect on depressive symptoms, g=-1.24 (95%CI -1.83 to -0.65, p< 0.001). Heterogeneity was small to moderate and NS. **Favors exercise**

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
				<u>Exercise vs Usual Care</u> Four studies of 180 participants showed a moderate and significant effect on depressive symptoms, g=-0.48 (95%CI -0.80 to 0.16, p< 0.001). Heterogeneity was NS. **Favors exercise** **<u>Exercise vs Psychotherapy</u>** Three studies of 79 participants showed a small, non-significant reduction of depressive symptoms g=-0.22 (95% CI = -0.65 to 0.21, p = 0.31. There was no substantial heterogeneity. **<u>Exercise vs. Anti-depressant Pharmacotherapy</u>** Three studies of 236 participants found no significant effect of exercise compared to medication, g = -0.08 (95%CI = -1.10 to 0.11, p=0.11) **<u>Exercise and Anti-depressants vs. medication alone</u>** Four studies of 188 participants yielded findings that were NS. **<u>Exercise vs. Control (Blinded Subgroup)</u>**

Study Details	Search Strategy/Evidence Base	Patients	Interventions/Comparators	Results
				In 13 studies with 377 participants, unblinded studies showed a large & significant effect in favor of exercise, g= -0.91(95%CI= -1.22 to -0.61, p<0.001), while the effect for the 10 studies with blinded outcome (600 participants) was moderate & significant, g=-0.40 (95% CI= - 0.69 to –0.11, p=0.01 showing a moderate and significant effect. **Favors exercise** **<u>Exercise versus control (with allocation concealment, ITT, and blinded outcome)</u>** Six studies with 461 participants included all 3 of these characteristics. Meta-analysis found that reduction in the depressive symptoms in these studies was small and NS. G=-0.26 (95%CI = -0.61 to 0.08, p=0.14). Heterogeneity was high and significant (I^2 = 68.53, p=0.007)

Table 4. Systematic Review Risk of Bias AMSTAR Checklist Table on Exercise to Treat MDD

Question	Kvam et al., 2016
Did the research questions and inclusion criteria for the review include the components of PICO?	Yes
Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	No
Did the review authors explain their selection of the study designs for inclusion in the review?	Yes
Did the review authors use a comprehensive literature search strategy?	Yes
Did the review authors perform study selection in duplicate?	Yes
Did the review authors perform data extraction in duplicate?	Yes
Did the review authors provide a list of excluded studies and justify the exclusions?	Yes
Did the review authors describe the included studies in adequate detail?	No
Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	Yes
Did the review authors report on the sources of funding for the studies included in the review?	No
If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results? RCTs?	Yes
If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	No
Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	Yes
Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	Yes
If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	Yes
Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	Yes
Overall Quality	**Moderate**

Table 5. AMSTAR Rating of Overall Confidence in Results of the Review

Category	Definition
High	*No or one non-critical weakness*: the systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest.
Moderate	*More than one non-critical weakness*: the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review.
Low or Very Low	*One or more critical flaw(s) with or without non-critical weaknesses*: the systematic review has one or more critical flaws and may not provide an accurate and comprehensive summary of the available studies that address the question of interest.

AMSTAR checklist, go to https://amstar.ca/Amstar_Checklist.php

Table 6. Evidence Table for RCTs on Exercise to Treat MDD

Study Details	Study Population	Treatment	Results	Conclusion/Limitations
Reference: Abdollahi 2017 **Purpose:** Determine the efficacy of exercise as an adjunct to CBT for suicidal ideation and depression **Setting:** Tehran, Iran from 2 psychology clinics **Funding source:** Not disclosed	**Number of patients:** 70 outpatients **Inclusion criteria:** Sedentary individuals with Mild to Moderate MDD **Exclusion criteria:** medical problems, stroke within 1 year, CVD, neuropathy, Parkinson's ds., vestibular disorders, severe visual problems, and severe depression (BDI-II ≥30), bipolar disorder, schizoaffective disorder or pregnancy **Pt. baseline characteristics:** mean age: 49.67(48;51), 37 men and 33 women; 51% married	**Intervention:** CBT+exercise: n = 35 **Control:** CBT only: n= 35 **Outcomes:** BDI-II; BSSI; BIADL**: F/U:** 12 weeks	**Depression:** r^2 =0.84, (95% CI :1.43. 2.83 Greater improvement in depression in the CBT/exercise compared to CBT alone when the model controls for suicidal ideation and ADL **Suicidal Ideation:** r^2= 0.82, (95% CI:1.16, 1.92) Greater improvement in suicidal ideation in the CBT/exercise compared to CBT alone when the model controls for depression ideation and ADL	**Conclusion:** Combined CBT and exercise led to greater improvements in suicidal ideation and depression when compared to CBT alone. **Limitations:** No controls No F/U after 12 weeks of treatment; long term results are unknown **Study ROB:** High **Author conflict:** None disclosed
Reference: Belvederi -Murri et al. 2015 **Purpose:** to determine whether exercise as an adjunct to sertraline leads to better outcomes in late life MDD	**Number of patients:** 121 **Inclusion criteria:** 65 – 85 years old with **MDD; HAMD** ≥18; sedentary **Exclusion criteria:** other Axis I dx; alcohol misuse; cognitive	**Intervention:** Group exercise sessions Sertraline plus non-progressive exercise (S+NPE) n = 37: Three 60-minute session/wk. for 24 weeks designed to improve strength, balance, respiration and	(Sertraline plus progressive aerobic [S+PAE] vs sertraline plus lower intensity non-progressive aerobic exercise [S+NPE]) HSRD 4 wks.: Both exercise	**Conclusion:** Greater decreases in remission were observed in the two exercise groups when compared to the pharmacotherapy alone group. **Limitations:** Assessor single-blind only No active comparator for social interaction or health education No measure of fidelity to protocol

Setting: Italy - 4 regional liaison programs between MH and primary care **Funding source:** Emilia Romagna Regional University Programme	impairment; severe or unstable physical illness. **Patient baseline characteristics:** **Mean age:** 75 years **Gender:** majority women **Education:** Majority elementary or less **Mean BMI:** 25 – 26 **Median HSRD: 18-19** **Age at onset of Depression:** 49 – 50 years **Treated with an antidepressant (lifetime):** 25 – 30 %	motor coordination and comprised of mat work and equipment. Exercise intensity was monitored by heart rate at up to 70% max. Group sessions were held with 4 - 6 participants. Sertraline plus progressive aerobic exercise n = 42: Exercise overlapped the S+NPE group but focused mainly on exercise to improve cardiopulmonary condition. **Control:** Sertraline only n = 42 **Outcomes:** rate and time to remission from depression over 24 weeks. Secondary outcomes depressions severity (HSRD); global improvement in depression (CGI) **F/U:** 24-week trial; no longer term FU after trial period.	groups showed improvement over sertraline alone. Pt rated "much improved": S-PAE – 71%, S+NPE – 60%, S – 43%, $X^2 = 7.09$. d.f. = 2, p=0.03 Remission rates at 24 wks. were higher in the S+PAE and S+NPE groups than the sertraline group. 81%, 73%, and 45%, respectively, p=0.001	Imbalance in patient baseline characteristics Not generalizable to younger population **Study RoB:** High **Author conflict:** None disclosed
Reference: Combs et al. 2014 **Purpose:** Examine the effects of exercise and sertraline on disordered sleep in MDD. Secondary analysis of the SMILE	**Number of patients:** 202 **Inclusion criteria:** DSM-IV MDD **Exclusion criteria:** At entry, any pt. already taking medication for MDD or sleep disturbance or	**Intervention:** **Group 1:** supervised AE for 3 45-minute sessions/wk. **Group 2:** home-based AE identical to Group 1 with exercise log. Home AE begun after training with an exercise	All four groups showed improvements in sleep on the ITT analysis at 16 weeks. (p=0.004). Importantly, active treatments showed more improvement	**Conclusions:** All 4 groups showed improvement in sleep, with active treatments showing more improvement than placebo. Improvements in sleep were comparable for exercise and sertraline. No significant difference found in early, middle, and late insomnia. At 1-year posttreatment, residual insomnia symptoms were associated with elevated depressive symptoms. **Limitations:** Secondary analysis of SMILE-II study. (Included in Kvam et al.). Relied on self-report and

study of 202 patients with MDD **Setting:** U.S. **Funding source:** Not disclosed	psychotherapy or exercising was excluded **Patient baseline characteristics:** 55% had early sleep complaints; 42% had middle sleep complaints; 45% had late sleep complaints. Participants available for 1 yr. FU did not differ in age (p=0.90), sex (p=0.06), ethnicity (p=0.52), or baseline depression severity (p=0.39)	physiologist and FU at 1 month & 2 months. **Control:** **Group 3:** sertraline titrated from 50 - 200 mg. Up to 4 doses of zolpidem allowed during the 4 wk study. **Group 4:** placebo control **Outcomes:** **Insomnia Symptoms from secondary analysis** MDD: diagnosed by SCID at baseline, 16 wk., and 1-year FU. Insomnia: HAMD: 3 item questions on insomnia.	than placebo (p=0.867) Improvements in sleep were comparable for exercise and sertraline. (p=0.841) and the pattern was unchanged when early responders were eliminated. (p=0.690). No differences were noted in early, middle and late insomnia. (p=0.071, 0.147, 0.871, respectively) One year: residual insomnia symptoms after treatment was associated with elevated depressive symptoms at 1-year (beta = -0.22, p=0.005). Residual insomnia was associated with risk of MDD diagnosis at 1 year (OR – 1.44, CI 1.11-1.87; p=0.006). Women (OR –	data from a 3-item subset of the HAMD rather than standard, validated insomnia questionnaires. Sertraline and placebo groups were double blind while the subjects of the exercise group were not blinded. **Study RoB:** High **Author Conflict:** None disclosed

			3.11; ci 1.15-8.39; p=0.026) and physical activity levels were also associated with MDD dx at 1 year (OR 0.97; CI 0.94-0.99; p=0.013)	
Reference: Doose 2015 **Purpose:** to examine the effect of exercise for patients with unipolar depression **Setting:** German outpatients from population-based recruitment **Funding source:** Robert Enke Foundation	**Number of patients:** 46 **Inclusion criteria:** 18 – 65 yr. ICD10 MDD **Exclusion criteria:** Psychosis, bipolar disorder, or other serious medical or co-occurring MH conditions, current psychopharmacologic treatment, or condition prohibiting exercise. **Patient baseline characteristics:** Mean age: 48 yr. Gender: women 63% Recurrent MDD in all but 3 pt.; mean years since 1st episode: 9 Ongoing pharmacotherapy: 58.7% Ongoing psychotherapy:58.7% HSRD-17: 14.21±3.08 BDI-II: 26.02±9.04 No differences in baseline characteristics	**Intervention:** supervised exercise 3x/wk. for 60 minutes for 8 wk. **Control:** wait list Depression symptoms measured by HSRD-17 and BDI-II	**Outcomes:** HSRD-17 was assessed by an unblinded clinician **ANCOVA HSDR-17:** 8.24 (CI -11.25-5.02; p=<0.0001) Clinical significance: 19 participants "recovered; 4 were rated as" improved" **BDI-II:** 2.04(CI -0.21-4.30; p=0.075) Clinical significance: Both cohorts rated as unchanged FU: 8 weeks	**Conclusions:** Exercise as an adjunct to TAU demonstrated a larger reduction in depressive symptoms than TAU alone or WL. **Limitations:** 24% attrition; Only 58% of sessions were attended; HSDR-17 was assessed by an unblinded researcher; BDI-II is a self-assessment; 8-week study with no long-term FU **Study RoB:** High **Author conflict:** None Disclosed

	between intervention and control			
Reference: Kerling et al. **Purpose:** examine the effectiveness of exercise as an adjunct to inpatient MDD treatment **Setting:** German hospital **Funding source:** Not funded	**Number of patients:** 42 inpatients Prospective randomized control trial of exercise (n=22) or TAU (n=20) **Inclusion criteria**: Adults > 18 years hospitalized in specialized psychotherapy ward with a diagnosis of moderate –severe MDD by DSM VI as assessed by SCID **Exclusion criteria:** severe medical condition or co-occurring MH or SUD diagnosis. **Patient baseline characteristics:** No significant differences in age, BMI, smoking, exercise before admission, and gender. AE group reported more alcohol consumption, higher waist circumference and diastolic BP.	All patients received CBT 90% participation in exercise sessions and no dropouts Anti-depressant medication was received by 17/22 (77%) of exercise intervention participants and 15/20 (75%) of TAU. **Intervention:** supervised exercise training of 3 AE sessions/wk. for 45 minutes at moderate intensity. 25min. bicycle ergometer; 20 min. cross trainer, stepper, treadmill, or rowing. **Control:** CBT +/- pharmacotherapy	**Outcomes:** depression severity was reduced with large effect sizes in both groups **MADRS sum score:** no significant difference **MADRS responders (50% reduction):** 64% AE vs 30% of TAU pts.; p=0.037 **BDI-II:** no significant difference between groups 77% of AE group and 75% of TAU were discharged on anti-depressants **F/U:** 6 weeks	**Conclusions:** While both groups experienced a decrease in depressive symptoms, the change was not statistically significant. More patients in the exercise group were classified as responders, defined as at least 50% reduction in MADRS score compared to TAU alone. **Limitations:** Hospitalized pt. population; Short duration without Long-term FU **Study RoB:** High **Author conflict:** None disclosed
Reference: Legrand 2015 Pur**pose:** Examine the effectiveness of short (10 day) course of individual	**Number of patients:** 35; exercise vs. TAU for MDD **Inclusion criteria:** Adults with a DSM-IV diagnosis of MDD;	**Intervention:** 30 minutes of brisk walking or jogging outdoors on 10 consecutive days (n=14)	**Self-report BDI_II** ANOVA: AE VS TAU: SD + -17.22vs. -6.41, p=0.012	**Conclusions:** While the ITT analysis showed significant improvements in self-reported depression for both the aerobic exercise and stretching groups, a large effect size at comparing pre- and post- depression changes was in favor of aerobic exercise. No

AE/endurance training to inpatient MDD treatment **Setting:** hospitalized patients in France **Funding source:** Not disclosed	antidepressant therapy for, 2 wk.; BDI-II score of ≥29; able to run or walk briskly **Exclusion criteria:** Not able to read French; medical contraindication; beta-blockers or another MDD treatment (e.g. ECT) **Patient baseline characteristics:** Mean age: 45 Women 71.4% No other significant demographic differences in randomized groups	**Control:** 1) Stretching for 30 minutes on 10 consecutive days (n=11) 2)no intervention/TAU = medication (n=10) **Outcomes:** Self-reported depression symptoms using BDI-II **F/U: 10 days**	AE vs. ST: -17.22 vs – 9.39, p=0.082 Effect sizes large Cohen's d =-1.39 ST vs. TAU No difference, p=1.000	significant changes in depressive symptoms were found in the control group of no intervention. **Limitations:** Small study of short duration in hospital patients Only 10 days of exercise 10/48 eligible patients refused participation. 7.1% attrition in AE 18.2% attrition in ST And 1 pt. in TAU failed to complete at least 8 measurements **Study RoB:** High **Author conflict:** None disclosed
Reference: Schuch et al. 2015 **Purpose:** to evaluate adjunct exercise in the treatment of severe MDD **Setting:** Hospitalized patients in Brazil **Funding source:** Fundo de Incentivo a Pesquisa e Eventos do Hospital de Clinicas de Porto Allegre	**Number of patients:** 50 inpatients **Inclusion criteria:** Adults, 18 – 60 with DSM-VI diagnosis of severe MDD, HAMD score of >25, **Exclusion criteria:** beta blocker treatment, co-occurring MH or SUD, cardiac risk factors or any medical condition that limits exercise **Patient baseline characteristics:** Mean age: 40.3 years, weight = 66.85 kg, BMI = 25.15, 58% non-smokers.	**Intervention:** Individual AE plus TAU n=25 Exercise 3x/week at 373 kcal/session on stationary bike, treadmill, or transport machine. Control: TAU = anti-depressant RX or ECT n=25 No psychotherapy. The distribution of TAU treatments was not different between groups. **Outcomes:**	At discharge difference of 3.70 (CI 6.21 -1.19, p=0.005) No significant difference between the AE and TAU groups in remission, p=0.248 or response, p=0.114. **WHOQoL-BREF:** Psychological QoL: 12.99 point, (CI 1.68-24.29), p=0.025 at second wk. and 19.10, CI 9.58-	**Conclusions:** No significant difference in remission or response rates was found between the exercise + TAU group and the TAU alone group. However, significant differences in QoL were found in favor of the exercise group. **Limitations:** Despite the low attrition, the acceptability was low with 52.8% refusing to participate Patients were not blinded. Small sample size for remission but NNT =6.2. **Study ROB:** High **Author Conflict:** None reported

	No difference in demographic characteristics of gender, age, weight, previous tobacco use, or family history of depression.	Reduction in depression symptoms as measured by HAMD HAMD – 17: score of < 7 is a "remission"; decrease of 50% is "response" WHOQoL-BREF: quality of life	28.62, **p=0.01 at discharge.** Three dropouts; 2 in the exercise group (8%) and 1 in the TAU (4%) group **F/U:** Mean duration of hospitalization was 23.6 (9.0) for the AE pts. TAU and 21.32 (8.2) for the TAU groups. **There was no difference in the length of stay (F=0.68; p=0.41)**	
Reference: Siquiera et l. 2016 **Purpose:** to evaluate the efficacy of exercise as an adjunct to sertraline **Setting:** Brazil outpatients **Funding source:** Sao Paulo Research Foundation (FAPESP)	**Number of patients:** 57 **Inclusion criteria:** Adults age 18-55 yr. with DSM VI dx of MDD; drug-free for at least 5 wk. prior to study enrollment. **Exclusion criteria:** Other co-occurring MH or SUD, medical conditions that limit exercise **Patient baseline characteristics:** **Mean age:** 38.82 ± 10.7, Women: n=41 Men: n=16	**Intervention:** Sertraline plus adjunct AE 4x/wk for 4wk. n=29 **Control:** Sertraline plus no activity, n= 28 **Outcomes:** HAMD -17: clinician measure of depression BDI-II: Self-report of depression symptoms	**HAMD -17**: No significant differences, p=0.84 BDI II: Both improved but there was no significant difference, p=0.35. AE intervention group had lower prescribed dosages of sertraline. FU: 28 days	**Conclusions:** No significant difference in depression was found between the exercise + sertraline group and the sertraline alone group. **Limitations:** 29.8 % attrition; Single blind of assessor only; Short duration and no long-term FU. **Study RoB:** High **Author conflict:** Not disclosed

Reference: Verrusio et al. 2014 **Purpose**: Evaluate the effectiveness of AE/music listening on mild-moderate MDD in the elderly **Setting:** Italy outpatients **Funding source:** Unknown	**Number of patients:** 24 **Inclusion criteria:** Adults with DSMVI diagnosis of mild – moderate MDD. **Exclusion criteria:** Conditions that limit exercise **Patient baseline characteristics:** Mean age: 75.5 Women: 13; Men: 11	**Intervention:** 2 one-hour sessions of exercise /wk. while listening to music. AE was stationary bike or treadmill. **Control:** paroxetine No anti-depressants during the study period. **Outcomes:** Geriatric Depression Scale (GDS): Hamilton Anxiety Scale: **FU:** 24 wk. all participants; 28 wk for AE/music group No dropouts	**Geriatric Depression Scale (GDS):** Wk 12: F=1.08, p=0.31 Wk 24: F= 10.44 p=0.01 **Hamilton Anxiety Scale:** Wk 12: F= 5.14, p=0.03 Wk 24: F=11.92, p=0.00	**Conclusions:** Those in the exercise and music group had greater improvements in self-reported depression, while the pharmacotherapy group had greater improvements in anxiety. **Limitations:** Small pilot study; Single blind **Study RoB:** High **Author conflict:** Not disclosed

AE: aerobic Exercise; BDI-II: BIADL: Barthel Index of Activities of Daily Living; Beck Depression Inventory-II; BMI: Body Mass Index; BSSI: Beck Scale for Suicidal Ideation; CBT: Cognitive Behavior Therapy; DSM – IV: Diagnostic Statistical Manual Version IV; GDS:Geriatric Depression Scale; HAMD: Hamilton Assessment of Major Depression; HAS: Hamilton Anxiety Scale; HSRD – 17:Hamilton Rating Scale for Depression – 17 questions; ICD 10: International Classification of Diseases Version 10; MDD: major depressive disorder; NS: not statistically significant RoB: risk of bias; SMD: standarized mean difference; TAU: treatment as usual; WHOQOL: World Health Quality of Life Brief Form;

Table 7. Cochrane Risk of Bias 2.0 Tool for RCTs on Exercise to Treat Major Depressive Disorder

Reference	Abdollahi et al., 2017	Belvederi-Murri et al. 2015	Combs et al., 2014	Doose et al., 2015	Kerling et al., 2015
• Was the allocation sequence generated adequately (e.g., random number table, computer-generated randomization)?	Yes	Yes	Yes	Yes	Yes
• Was the allocation of treatment adequately concealed (e.g., pharmacy-controlled randomization, concealed envelopes)?	NI	Yes	Yes	Yes	Yes
• Did baseline difference between study groups suggest a problem with randomization?	No	No	No	No	No
Overall RoB for Randomization Process	**Some concern**	**Low**	**Low**	**Low**	**Low**
Deviation from Intended Intervention (Effect of Assignment)					
• Were participants aware of their assigned intervention during the trial?	Yes	Yes	Yes	Yes	Yes
• Were providers and people delivering treatment aware of assigned intervention during trial?	Yes	No	No	Yes	PY
• Were there deviations from the intended intervention that arose because of the experimental context?	No	Yes	PY	Yes	No
• Were these deviations from intended intervention balanced between groups?	NA	No	PN	No	Yes
• Were these deviations likely to have affected the outcome?	NA	No	Yes	Yes	NA
• Was an appropriate analysis used to estimate the effect of assignment to intervention?	No	Yes	Yes	Yes	Yes
Overall RoB of Effect of Assignment	**Some concern**	**High**	**High**	**High**	**High**
Missing Outcome Data					
• Were data for this outcome available for all, or nearly all, participants randomized?	No	Yes	PN	No	Yes
• Is there evidence that result was not biased by missing outcome data?	No	Yes	PY	No	Yes
• Could missingness in the outcome depend on its true value?	PN	Yes	Yes	Yes	No
• Do the proportions of missing outcome data differ between intervention groups?	PN	No	No	No	No
• Is it likely that missingness in the outcome depended on its true value?	PN	PY	PY	No	No

Reference	Abdollahi et al., 2017	Belvederi-Murri et al. 2015	Combs et al., 2014	Doose et al., 2015	Kerling et al., 2015
Overall RoB of Missing Data	**High**	**Some Concern**	**High**	**High**	**Low**
Measurement of the Outcome					
• Was the method of measuring the outcome inappropriate?	No	No	Yes	No	No
• Could measurement or ascertainment of the outcome have differed between intervention groups?	No	No	PY	Yes	PY
• Were outcome assessors aware of the intervention received by study participants?	No	No	No	Yes	Yes
• Could assessment of the outcome have been influenced by knowledge of intervention received?	Yes	Yes	Yes	Yes	Yes
• Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	No	PY	Yes	PY	PY
Overall RoB of Measurement of Outcome	**High**	**High**	**High**	**High**	**High**
Selection of Reported Results					
• Was the trial analyzed in accordance with a pre-specified plan that was finalized before unblinded outcome data were available for analysis?	Yes	Yes	Yes	Yes	Yes
Overall RoB of Reported Results	**Low**	**Low**	**Low**	**Low**	**Low**
Overall Study RoB	**High**	**High**	**High**	**High**	**High**

Table 7. Cochrane Risk of Bias 2.0 Tool for RCTs on Exercise to Treat Major Depressive Disorder (continued)

Reference	Legrand et al., 2016	Schuch et al., 2015	Siqueira et al., 2016	Verrusio et al., 2014
• Was the allocation sequence generated adequately (e.g., random number table, computer-generated randomization)?	Yes	Yes	Yes	Yes
• Was the allocation of treatment adequately concealed (e.g., pharmacy-controlled randomization, concealed envelopes)?	Yes	Yes	Yes	Yes
• Did baseline difference between study groups suggest a problem with randomization?	No	No	No	PN
Overall RoB for Randomization Process	**Low**	**Low**	**Low**	**Low**

Reference	Legrand et al., 2016	Schuch et al., 2015	Siqueira et al., 2016	Verrusio et al., 2014
Deviation from Intended Intervention (Effect of Assignment)				
• Were participants aware of their assigned intervention during the trial?	Yes	Yes	Yes	Yes
• Were providers and people delivering treatment aware of assigned intervention during trial?	PY	PY	No	PY
• Were there deviations from the intended intervention that arose because of the experimental context?	No	Yes	Yes	PY
• Were these deviations from intended intervention balanced between groups?	Yes	No	No	PN
• Were these deviations likely to have affected the outcome?	Yes	Yes	Yes	Yes
• Was an appropriate analysis used to estimate the effect of assignment to intervention?	No	Yes	Yes	NA
Overall RoB of Effect of Assignment	**High**	**High**	**High**	**High**
Missing Outcome Data				
• Were data for this outcome available for all, or nearly all, participants randomized?	No	Yes	No	Yes
• Is there evidence that result was not biased by missing outcome data?	Yes	Yes	Yes	Yes
• Could missingness in the outcome depend on its true value?	Yes	Yes	Yes	NA
• Do the proportions of missing outcome data differ between intervention groups?	No	Yes	Yes	NA
• Is it likely that missingness in the outcome depended on its true value?	No	Yes	PY	NA
Overall RoB of Missing Data	**Some Concern**	**High**	**High**	**Low**
Measurement of the Outcome				
• Was the method of measuring the outcome inappropriate?	No	No	No	Yes
• Could measurement or ascertainment of the outcome have differed between intervention groups?	Yes	Yes	Yes	Yes
• Were outcome assessors aware of the intervention received by study participants?	Yes	No	No	Yes
• Could assessment of the outcome have been influenced by knowledge of intervention received?	Yes	Yes	Yes	Yes
• Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	PY	PY	PY

Reference	Legrand et al., 2016	Schuch et al., 2015	Siqueira et al., 2016	Verrusio et al., 2014
Overall RoB of Measurement of Outcome	**High**	**High**	**High**	**High**
Selection of Reported Results				
• Was the trial analyzed in accordance with a pre-specified plan that was finalized before unblinded outcome data were available for analysis?	Yes	Yes	Yes	Yes
Overall RoB of Reported Results	**Low**	**Low**	**Low**	**Low**
Overall Study RoB	**High**	**High**	**High**	**High**

Table 8. Cochrane Risk of Bias 2.0 Overall Risk of Bias Judgement

Category	Definition
Low risk of bias	The study is judged to be at low risk of bias for all domains for this result.
Some concerns	The study is judged to be at some concerns in at least one domain for this result.
High risk of bias	The study is judged to be at high risk of bias in at least one domain for this result. OR The study is judged to have some concerns for multiple domains in a way that substantially lowers confidence in the result.

References

Abdollahi, A., LeBouthillier, D. M., Najafi, M., Asmundson, G. J., Hosseinian, S., Shahidi, S., … Jalili, M. (2017). Effect of exercise augmentation of cognitive behavioural therapy for the treatment of suicidal ideation and depression. *Journal of Affective Disorders, 219*, 58–63. https://doi.org/http://dx.doi.org/10.1016/j.jad.2017.05.012

Belvederi Murri, M., Amore, M., Menchetti, M., Toni, G., Neviani, F., Cerri, M., … Efficacy of Exercise for Depression in Seniors Study, G. (2015). Physical exercise for late-life major depression. *British Journal of Psychiatry*, *207*(3), 235–242. https://doi.org/https://dx.doi.org/10.1192/bjp.bp.114.150516

Combs, K., Smith, P. J., Sherwood, A., Hoffman, B., Carney, R. M., Freedland, K., … Blumenthal, J. A. (2014). Impact of sleep complaints and depression outcomes among participants in the standard medical intervention and long-term exercise study of exercise and pharmacotherapy for depression. *Journal of Nervous and Mental Disease*, *202*(2), 167–171. https://doi.org/http://dx.doi.org/10.1097/NMD.0000000000000085

Doose, M., Ziegenbein, M., Hoos, O., Reim, D., Stengert, W., Hoffer, N., … Sieberer, M. (2015). Self-selected intensity exercise in the treatment of major depression: A pragmatic RCT. *International Journal of Psychiatry in Clinical Practice*, *19*(4), 266–275. https://doi.org/https://dx.doi.org/10.3109/13651501.2015.1082599

Kerling, A., Kuck, M., Tegtbur, U., Grams, L., Weber-Spickschen, S., Hanke, A., … Kahl, K. (2017). Exercise increases serum brain-derived neurotrophic factor in patients with major depressive disorder. *Journal of Affective Disorders*, *215*, 152–155. https://doi.org/http://dx.doi.org/10.1016/j.jad.2017.03.034

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

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

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

Siqueira, C. C., Valiengo, L. L., Carvalho, A. F., Santos-Silva, P. R., Missio, G., de Sousa, R. T., … Machado-Vieira, R. (2016). Antidepressant efficacy of adjunctive aerobic activity and associated biomarkers in Major Depression: A 4-week, randomized, single-blind, controlled clinical trial. *PLoS ONE*, *11*, e0154195. https://doi.org/https://dx.doi.org/10.1371/journal.pone.0154195

Verrusio, W., Andreozzi, P., Marigliano, B., Renzi, A., Gianturco, V., Pecci, M., …Gueli, N. (2014). Exercise training and music therapy in elderly with depressive syndrome: A pilot study. *Complementary Therapies in Medicine, 22*, 614-620.

Appendix A

Inclusion Criteria:

- **Publications type:** Systematic reviews (SRs) and randomized controlled clinical trials (RCTs) published in English language in peer reviewed journals.
- **Search date:** 01/01/2008 to present
- **Population:** Adults 18 years or older meeting diagnostic criteria for MDD
- **Intervention (s):**
    - Complementary and integrative health (CIH) and other non-pharmacologic treatments: music therapy; equine therapy; training and caring for service dogs; yoga therapy; tai chi; acupuncture therapy; meditation therapy; outdoor sports therapy; hyperbaric oxygen therapy; accelerated resolution therapy; art therapy; magnetic stimulation therapy; massage; healing touch; therapeutic touch; cannabinoids; chiropractic care
    - Pharmacological treatments: SSRIs (fluoxetine, paroxetine, escitalopram, citalopram, vilazodone, and vortioxetine); SNRIs (duloxetine, venlafaxine, levomilnacipram, and desvenlafaxine); tetracyclic antidepressants (mirtazapine); NDRI (bupropion); ketamine
    - Psychological treatments: Acceptance and Commitment Therapy; behavioral therapy; behavioral activation; CBT; computer-based CBT; interpersonal therapy; MBCT; and problem-solving therapy;
- **Outcomes:** improvement in global MDD severity, adverse events; loss of diagnosis; remission; self-reported MDD symptom improvement; comorbid symptoms; quality of life; functional status; patient satisfaction; anxiety; insomnia; and pain
- **Timing:** no minimum follow-up
- **Setting(s):** primary care; specialty care; general mental health care

Exclusion Criteria:

- **Wrong publication type**: narrative review article, case reports editorial, commentary, protocol of randomized trial without results, any article without original data, abstract alone.
- **Wrong study design:** Observational study (for example, cohort study, case control study, cross-sectional study); treatment study without randomization, randomized study with less than 20 patients (10 per study group).
- **Wrong population**: animal studies, children or adolescents less than 18 years of age (studies must have enrolled a patient population in which at least 80% of patients were diagnosed with MDD.
- **Wrong language:** Study in language other than English.
- **Wrong or no intervention:** CIH or other non-pharmacologic treatments other than those listed in inclusion criteria; medications other than those listed in inclusion criteria; psychological treatments other than those listed in inclusion criteria
- **Wrong comparator:** CIH or other non-pharmacologic treatments other than those listed in inclusion criteria; medications other than those listed in inclusion criteria; psychological treatments other than those listed in inclusion criteria

- **Wrong outcome(s):** Any study that does not have at least one of the included outcomes of interest. Any subjective outcome (e.g. symptoms; quality of life) not measured using a validated instrument.

Appendix B

Table 1. Studies Excluded at Full-text Level

Authors	Reason for Exclusion
Acupuncture	
Li, 2018	Wrong population
Wen, 2018	Wrong population
Asher, 2017	More recent/comprehensive SR available
Dong, 2017	More recent/comprehensive SR available
Gartlehner, 2017	More recent/comprehensive SR available
Asher, 2016	More recent/comprehensive SR available
Chung, 2016	Wrong population
Gartlehner, 2016	More recent/comprehensive SR available
Chung, 2015	Included in Smith, 2018
Fan, 2015	Duplicate publication of Fan, 2013
Gartlehner, 2015	More recent/comprehensive SR available
Liu, 2015	Wrong outcome
Chen, 2014	Wrong outcome
Hopton, 2014	Wrong population
Wang, 2014	Included in Smith, 2018
MacPherson, 2013	Included in Smith, 2018
Qu, 2013	Included in Smith, 2018
Sun, 2013	Included in Smith, 2018
Wang, 2013	Wrong outcome
Zhang, 2013	Wrong outcome
Ma, 2012	Wrong comparator
Zhang, 2012	Included in Smith, 2018
Andreescu, 2011	Included in Smith, 2018
Duan, 2011	Included in Smith, 2018
Feng, 2011	Included in Smith, 2018
Yeung, 2011	Included in Smith, 2018
Manber, 2010	Wrong population
Smith, 2010	More recent/comprehensive SR available
Duan, 2009	Duplicate publication of Duan, 2011
Zhang, 2009	Included in Smith, 2018
Whiting, 2008	Wrong population
Art Therapy	
Dunphy, 2018	Wrong study design
Nan, 2017	Wrong comparator
Exercise	

Authors	Reason for Exclusion
Balchin et al., 2016	Wrong comparator
Bridle et al., 2012	More recent/comprehensive SR
Carneiro et al., 2015	Wrong population
Carta et al., 2008	Wrong outcomes
Catalan-Matamoros et al., 2016	More recent/comprehensive SR
Chu et al., 2009	Wrong population
Cooney et al., 2013	More recent/comprehensive SR
Danielsson et al., 2014	Wrong comparator
de la Cerda et al., 2011	Wrong study design
Greer et al., 2016	Wrong comparator
Heinzel et al., 2015	More recent/comprehensive SR
Helgadottir et al., 2016	Wrong population
Holvast et al., 2017	More recent/comprehensive SR
Huang et al., 2015	Wrong population
Josefsson et al., 2014	More recent/comprehensive SR
Kerling et al., 2017	Wrong outcomes
Kerse et al., 2010	Wrong population
Krogh et al., 2011	More recent/comprehensive SR
Luttenberger et al., 2015	Wrong population
Martiny et al., 2015	Wrong intervention
Mead et al., 2009	More recent/comprehensive SR
Meyer et al., 2016	Wrong comparator
Meyer et al., 2016b	Wrong outcomes
Minghetti et al., 2018	Wrong comparator
Morres et al., 2019	More recent/comprehensive SR
Mura et al., 2013	More recent/comprehensive SR
Neunhauserer et al., 2013	Wrong population
Neviani et al., 2017	Wrong outcomes
Nguyen et al., 2014	Wrong population
Rahman et al., 2018	Wrong outcomes
Rethorst et al., 2017	Wrong comparator
Rethorst et al., 2013	Wrong comparator
Rethorst et al., 2010	Wrong outcomes
Rhyner et al., 2016	More recent/comprehensive SR
Rimer et al., 2012	More recent/comprehensive SR
Schuch et al., 2016	More recent/comprehensive SR
Seo et al., 2018	More recent/comprehensive SR
Sherwood et al., 2016	Wrong outcomes
Silveira et al., 2013	More recent/comprehensive SR

Authors	Reason for Exclusion
Stanton et al., 2014	More recent/comprehensive SR
Strom et al., 2013	Wrong intervention
Stubbs et al., 2016	More recent/comprehensive SR
Sun et al., 2018	More recent/comprehensive SR
Toups et al., 2017	Wrong outcomes
Trivedi et al., 2011	Wrong comparator
Williams et al., 2008	Wrong population
Yeh et al., 2015	Wrong outcomes
Zanetidou et al., 2017	Secondary analysis of included study
Massage	
Hou, 2010	Wrong population
Meditation	
Capobianco, 2018	Wrong intervention
Ionson, 2018	Wrong outcome
Rajagpol, 2018	Wrong population
Jain, 2015	Wrong intervention
Rentala, 2015	Wrong population
Prakhinkit, 2014	Wrong population
Lo, 2013	Wrong population
Chan, 2012	Wrong intervention
Yang, 2009	Wrong outcome
Butler, 2008	Wrong population
Tsang, 2008	Wrong intervention
Music Therapy	
Trimmer, 2018	Wrong intervention
Aalbers, 2017	Wrong population
Leubner, 2017	Wrong population
Zhao, 2016	Wrong population
Kumar, 2013	Wrong study design
Castillo-Perez, 2010	Wrong intervention
Tai Chi	
Liao, 2018	Wrong study design
Liu, 2015	Wrong population
Yin, 2014	Wrong population
Chi, 2013	Wrong population
Therapeutic Touch	
Klainin-Yobas, 2015	Wrong study design
Zhao, 2012	Wrong population
Transcranial Magnetic Stimulation (TMS)	

Authors	Reason for Exclusion
Kim, 2019	Wrong population
Maneeton, 2019	Abstract only
Blumberger, 2018	Wrong comparator
Fava, 2018	Wrong intervention
Fitzgerald, 2018	Wrong comparator
Haesebaert, 2018	Wrong intervention
Kaster, 2018	Wrong intervention
Kavanaugh, 2018	Wrong population
Razza, 2018	Does not address key question
Brunoni, 2017	Wrong intervention
Carpenter, 2017	Wrong intervention
Kedzior, Muller, Gellersen, et al., 2017	Abstract only
Kedzior, Muller, Gerkensmeier, et al., 2017	Wrong intervention
Mutz, 2017	Abstract only
Theleritis, 2017	Wrong comparator
Wang, 2017	Wrong population
Yip, 2017	Wrong study design
Fitzgerald, 2016	Included in Brunoni, 2017
Philip, 2016	Does not address key question
Dell'Osso, 2015	Included in Brunoni, 2017
Kaur, 2015	Abstract only
Kedzior, 2015	Wrong intervention
Kreuzer, 2015	Included in Brunoni, 2017
Leggett, 2015	More recent/comprehensive review available
Levkovitz, 2015	Included in Brunoni, 2017
Prasser, 2015	Included in Brunoni, 2017
Serafini, 2015	Does not meet inclusion criteria for SR
Xie, 2015	Wrong comparator
Brunelin, 2014	Included in Brunoni, 2017
Gaynes, 2014	More recent/comprehensive review available
Leuchter, 2014	Included in Brunoni, 2017
Plewnia, 2014	Wrong intervention
Baeken, 2013	Included in Brunoni, 2017
Berlim, 2013	More recent/comprehensive review available
Chen, Zhou, et al., 2013	More recent/comprehensive review available
Chen, Chang, et al., 2013	Included in Brunoni, 2017
Fitzgerald, 2013	Included in Brunoni, 2017
Mantovani, 2013	Wrong population

Authors	Reason for Exclusion
Bakim, 2012	Included in Brunoni, 2017
Blumberger, Daniel, Mulsant, et al., 2012	Included in Brunoni, 2017
Blumberger, Daniel, Tran, et al., 2012	Included in Brunoni, 2017
Fitzgerald, 2012	Included in Brunoni, 2017
Galletly, 2012	Wrong comparator
Huang, 2012	Included in Brunoni, 2017
Aguirre, 2011	Included in Brunoni, 2017
Duan, 2011	Wrong outcomes
Fitzgerald, 2011	Included in Brunoni, 2017
Levkovitz, 2011	Wrong intervention
Nongpiur, 2011	Wrong intervention
Ray, 2011	Included in Brunoni, 2017
Avery, 2010	Included in Brunoni, 2017
George, 2010	Included in Brunoni, 2017
Janicak, 2010	Included in Brunoni, 2017
Martinot, 2010	Included in Brunoni, 2017
Pallanti, 2010	Included in Brunoni, 2017
Rossini, 2010	Included in Brunoni, 2017
Triggs, 2010	Included in Health Quality Ontario, 2016
Bares, 2009	Included in Brunoni, 2017
Carretero, 2009	Wrong population
Fitzgerald, 2009	Included in Brunoni, 2017
Levkovitz, 2009	Wrong intervention
Lisanby, 2009	Included in Brunoni, 2017
Schutter, 2009	Included in Brunoni, 2017
Avery, 2008	Included in Brunoni, 2017
Bretlau, 2008	Included in Brunoni, 2017
Mogg, 2008	Included in Health Quality Ontario, 2016
Yoga	
Tolahunase, 2018	Wrong outcomes
Prathikanti, 2017	Included in Vollbehr, 2018
Streeter, 2017	Wrong comparator
Falsafi, 2016	Included in Vollbehr, 2018
Naveen, 2016	Wrong outcomes
Schuver, 2016	Included in Vollbehr, 2018
Kinser, 2014	Included in Vollbehr, 2018
Gangadhar, 2013	Wrong population (diagnosis unclear)
Kinser, 2013	Included in Vollbehr, 2018
Mitchell, 2012	Included in Vollbehr, 2018

Authors	Reason for Exclusion
Shahidi, 2011	Wrong comparator
Butler, 2008	Included in Vollbehr, 2018

References

Aalbers, S., Fusar-Poli, L., Freeman, R. E., Spreen, M., Ket, J. C., Vink, A. C., … Gold, C. (2017). Music therapy for depression. *Cochrane Database of Systematic Reviews*, *11*, CD004517. https://doi.org/https://dx.doi.org/10.1002/14651858.CD004517.pub3

Aguirre, I., Carretero, B., Ibarra, O., Kuhalainen, J., Martinez, J., Ferrer, A., … Garcia-Toro, M. (2011). Age predicts low-frequency transcranial magnetic stimulation efficacy in major depression. *Journal of Affective Disorders, 130*(3), 466–469. https://doi.org/https://dx.doi.org/10.1016/j.jad.2010.10.038

Andreescu, C., Glick, R. M., Emeremni, C. A., Houck, P. R., & Mulsant, B. H. (2011). Acupuncture for the treatment of major depressive disorder: A randomized controlled trial. *Journal of Clinical Psychiatry, 72*(8), 1129–1135. https://doi.org/https://dx.doi.org/10.4088/JCP.10m06105

Asher, G., Gertlehner, G., Gaines, B., Amick, H., Morgan, L., Coker-Schwimmer, E., … Lohr, K. (2016). Comparative efficacy of antidepressants and complementary therapies for major depression: A systematic review and meta-analysis. *Journal of Alternative and Complementary Medicine, 22*(6), A103. https://doi.org/http://dx.doi.org/10.1089/acm.2016.29003.abstracts

Asher, G. N., Gartlehner, G., Gaynes, B. N., Amick, H. R., Forneris, C., Morgan, L. C., … Lohr, K. N. (2017). Comparative benefits and harms of complementary and alternative medicine therapies for initial treatment of major depressive disorder: Systematic review and meta-analysis. *Journal of Alternative & Complementary Medicine, 23*(12), 907–919. https://doi.org/https://dx.doi.org/10.1089/acm.2016.0261

Avery, D. H., George, M. S., Lisanby, S. H., McDonald, W. M., Durkalski, V., Pavlicova, M., … Sackeim, H. (2010). Daily left prefrontal TMS for major depression: A sham-controlled multi-site randomized trial. *Biological Psychiatry, 1*, 57S. https://doi.org/http://dx.doi.org/10.1016/j.biopsych.2010.03.007

Avery, D. H., Isenberg, K. E., Sampson, S. M., Janicak, P. G., Lisanby, S. H., Maixner, D. F., … George, M. S. (2008). Transcranial magnetic stimulation in the acute treatment of major depressive disorder: clinical response in an open-label extension trial. *Journal of Clinical Psychiatry, 69*(3), 441–451.

Baeken, C., Vanderhasselt, M. A., Remue, J., Herremans, S., Vanderbruggen, N., Zeeuws, D., … De Raedt, R. (2013). Intensive HF-rTMS treatment in refractory medication-resistant unipolar depressed patients. *Journal of Affective Disorders, 151*(2), 625–631. https://doi.org/https://dx.doi.org/10.1016/j.jad.2013.07.008

Bakim, B., Uzun, U. E., Karamustafalioglu, O., Ozcelik, B., Alpak, G., Tankaya, O., … Yavuz, B. G. (2012). The combination of antidepressant drug therapy and high-frequency repetitive transcranial magnetic stimulation in medication-resistant depression. *Bulletin of Clinical Psychopharmacology, 22*(3), 244–253. https://doi.org/http://dx.doi.org/10.5455/bcp.20120807092434

Balchin, R., Linde, J., Blackhurst, D., Rauch, H. L., & Schonbachler, G. (2016). Sweating away depression? The impact of intensive exercise on depression. *Journal of Affective Disorders*, *200*, 218–221. https://doi.org/https://dx.doi.org/10.1016/j.jad.2016.04.030

Bares, M., Kopecek, M., Novak, T., Stopkova, P., Sos, P., Kozeny, J., … Hoschl, C. (2009). Low frequency (1-Hz), right prefrontal repetitive transcranial magnetic stimulation (rTMS) compared with venlafaxine ER in the treatment of resistant depression: a double-blind, single-centre, randomized study. *Journal of Affective Disorders, 118*(1-3), 94–100. https://doi.org/https://dx.doi.org/10.1016/j.jad.2009.01.032

Berlim, M. T., Van Den Eynde, F., & Daskalakis, Z. J. (2013). High-frequency repetitive transcranial magnetic stimulation accelerates and enhances the clinical response to antidepressants in major depression: A meta-analysis of randomized, double-blind, and sham-controlled trials. *Journal of Clinical Psychiatry, 74*(2), e122–e129. https://doi.org/http://dx.doi.org/10.4088/JCP.12r07996

Blumberger, D. M., Mulsant, B. H., Fitzgerald, P. B., Rajji, T. K., Ravindran, A. V, Young, L., … Daskalakis, Z. J. (2012). A randomized double-blind sham-controlled comparison of unilateral and bilateral repetitive transcranial magnetic stimulation for treatment-resistant major depression. *The World Journal of Biological Psychiatry, 13*(6), 423–435. https://doi.org/http://dx.doi.org/10.3109/15622975.2011.579163

Blumberger, D. M., Tran, L. C., Fitzgerald, P. B., Hoy, K. E., & Daskalakis, Z. J. (2012). A randomized double-blind sham-controlled study of transcranial direct current stimulation for treatment-resistant major depression. *Frontiers in Psychology, 3*.

Blumberger, D. M., Vila-Rodriguez, F., Thorpe, K. E., Feffer, K., Noda, Y., Giacobbe, P., … Downar, J. (2018). Effectiveness of theta burst versus high-frequency repetitive transcranial magnetic stimulation in patients with depression (THREE-D): A randomised non-inferiority trial. *Lancet, 391*(10131), 1683–1692. https://doi.org/https://dx.doi.org/10.1016/S0140-6736(18)30295-2

Bretlau, L., Lunde, M., Lindberg, L., Unden, M., Dissing, S., & Bech, P. (2008). Repetitive transcranial magnetic stimulation (rTMS) in combination with escitalopram in patients with treatment-resistant major depression. A double-blind, randomised, sham-controlled trial. *Pharmacopsychiatry, 41*(2), 41–47. https://doi.org/http://dx.doi.org/10.1055/s-2007-993210

Bridle, C., Spanjers, K., Patel, S., Atherton, N. M., & Lamb, S. E. (2012). Effect of exercise on depression severity in older people: Systematic review and meta-analysis of randomised controlled trials. *British Journal of Psychiatry*, *201*(3), 180–185. https://doi.org/https://dx.doi.org/10.1192/bjp.bp.111.095174

Brunelin, J., Jalenques, I., Trojak, B., Attal, J., Szekely, D., Gay, A., … Poulet, E. (2014). The efficacy and safety of low frequency repetitive transcranial magnetic stimulation for treatment-resistant depression: the results from a large multicenter French RCT. *Brain Stimulation, 7*(6), 855–863. https://doi.org/https://dx.doi.org/10.1016/j.brs.2014.07.040

Brunoni, A., Moffa, A., Sampaio-Junior, B., Borrione, L., Moreno, M., Fernandes, R., … Bensenor, I. (2017). Trial of electrical direct-current therapy versus escitalopram for depression. *The New England Journal of Medicine, 376*(26), 2523–2533. https://doi.org/http://dx.doi.org/10.1056/NEJMoa1612999

Butler, L. D., Waelde, L. C., Hastings, T. A., Chen, X. H., Symons, B., Marshall, J., … Spiegel, D. (2008). Meditation with yoga, group therapy with hypnosis, and psychoeducation for long-term depressed mood: a randomized pilot trial. *Journal of Clinical Psychology*, *64*(7), 806–820. https://doi.org/https://dx.doi.org/10.1002/jclp.20496

Capobianco, L., Reeves, D., Morrison, A. P., & Wells, A. (2018). Group metacognitive therapy vs. mindfulness meditation therapy in a transdiagnostic patient sample: A randomised feasibility trial. *Psychiatry Research, 259*, 554–561. https://doi.org/https://dx.doi.org/10.1016/j.psychres.2017.11.045

Carneiro, L. S., Fonseca, A. M., Vieira-Coelho, M. A., Mota, M. P., & Vasconcelos-Raposo, J. (2015). Effects of structured exercise and pharmacotherapy vs. pharmacotherapy for adults with depressive symptoms: A randomized clinical trial. *Journal of Psychiatric Research*, *71*, 48–55. https://doi.org/https://dx.doi.org/10.1016/j.jpsychires.2015.09.007

Carpenter, L. L., Aaronson, S. T., Clarke, G. N., Holtzheimer, P. E., Johnson, C. W., McDonald, W. M., … Schneider, M. (2017). rTMS with a two-coil array: Safety and efficacy for treatment resistant major depressive disorder. *Brain Stimulation, 10*(5), 926–933. https://doi.org/http://dx.doi.org/10.1016/j.brs.2017.06.003

Carretero, B., Martin, M., Juan, A., Pradana, M., Martin, B., Carral, M., … Garcia-Toro, M. (2009). Low-frequency transcranial magnetic stimulation in patients with fibromyalgia and major depression. *Pain Medicine, 10*(4), 748–753. https://doi.org/http://dx.doi.org/10.1111/j.1526-4637.2009.00625.x

Carta, M. G., Hardoy, M. C., Pilu, A., Sorba, M., Floris, A. L., Mannu, F. A., … Salvi, M. (2008). Improving physical quality of life with group physical activity in the adjunctive treatment of major depressive disorder. *Clinical Practice and Epidemiology in Mental Health*, *4*(1).

Castillo-Perez, S., Gomez-Perez, V., Calvillo Velasco, M., Perez-Campos, E., & Mayoral, M.-A. (2010). Effects of music therapy on depression compared with psychotherapy. *The Arts in Psychotherapy, 37*(5), 387–390. https://doi.org/http://dx.doi.org/10.1016/j.aip.2010.07.001

Catalan-Matamoros, D., Gomez-Conesa, A., Stubbs, B., & Vancampfort, D. (2016). Exercise improves depressive symptoms in older adults: An umbrella review of systematic reviews and meta-analyses. *Psychiatry Research*, *244*, 202–209. https://doi.org/http://dx.doi.org/10.1016/j.psychres.2016.07.028

Chan, A. S., Wong, Q. Y., Sze, S. L., Kwong, P. P. K., Han, Y. M. Y., & Cheung, M. C. (2012). A Chinese Chan -based mind-body intervention improves sleep on patients with depression: a randomized controlled trial. *The Scientific World Journal*

Chen, J., Zhou, C., Wu, B., Wang, Y., Li, Q., Wei, Y., … Xie, P. (2013). Left versus right repetitive transcranial magnetic stimulation in treating major depression: A meta-analysis of randomised controlled trials. *Psychiatry Research, 210*(3), 1260–1264. https://doi.org/http://dx.doi.org/10.1016/j.psychres.2013.09.007

Chen, S. J., Chang, C. H., Tsai, H. C., Chen, S. T., & Lin, C. C. (2013). Superior antidepressant effect occurring 1 month after rTMS: add-on rTMS for subjects with medication-resistant depression. *Neuropsychiatric Disease and Treatment, 9*, 397–401.

Chen, J. Q., Lin, W. R., Wang, S. X., Wang, C. Q., Li, G. L., Qu, S. S., … Xiao, W. (2014). Acupuncture/electroacupuncture enhances antidepressant effect of seroxat: The symptom checklist-90 scores. *Neural Regeneration Research, 9*(2), 213–222.

Chi, I., Jordan-Marsh, M., Guo, M., Xie, B., & Bai, Z. (2013). Tai chi and reduction of depressive symptoms for older adults: a meta-analysis of randomized trials. *Geriatrics & Gerontology International, 13*(1), 3–12. https://doi.org/https://dx.doi.org/10.1111/j.1447-0594.2012.00882.x

Chu, I., Buckworth, J., Kirby, T. E., & Emery, C. F. (2009). Effect of exercise intensity on depressive symptoms in women. *Mental Health and Physical Activity*, *2*(1), 37–43. https://doi.org/http://dx.doi.org/10.1016/j.mhpa.2009.01.001

Chung, K.-F., Yeung, W.-F., Yu, Y.-M., Yung, K.-P., Zhang, S.-P., Zhang, Z.-J., … Chan, L.-W. (2015). Acupuncture for residual insomnia associated with major depressive disorder: A placebo- and sham-controlled, subject- and assessor-blind, randomized trial. *The Journal of Clinical Psychiatry, 76*(6), e752–e760. https://doi.org/http://dx.doi.org/10.4088/JCP.14m09124

Chung, K. F., Yeung, W. F., Zhang, S. P., Zhang, Z. J., Wong, M. T., Lee, W. K., & Chan, K. W. (2016). Acupuncture for persistent insomnia associated with major depressive disorder: A randomised controlled trial. *Hong Kong Medical Journal, 22*(2), S9-14.

Cooney, G. M., Dwan, K., Greig, C. A., Lawlor, D. A., Rimer, J., Waugh, F. R., … Mead, G. E. (2013). Exercise for depression. *Cochrane Database of Systematic Reviews*, (9), CD004366. https://doi.org/https://dx.doi.org/10.1002/14651858.CD004366.pub6

Danielsson, L., Papoulias, I., Petersson, E. L., Carlsson, J., & Waern, M. (2014). Exercise or basic body awareness therapy as add-on treatment for major depression: A controlled study. *Journal of Affective Disorders*, *168*, 98–106. https://doi.org/https://dx.doi.org/10.1016/j.jad.2014.06.049

de la Cerda, P., Cervello, E., Cocca, A., & Viviana, J. (2011). Effect of an aerobic training program as complementary therapy in patients with moderate depression. *Perceptual & Motor Skills*, *112*(3), 761–769.

Dell'Osso, B., Oldani, L., Camuri, G., Dobrea, C., Cremaschi, L., Benatti, B., … Altamura, A. C. (2015). Augmentative repetitive Transcranial Magnetic Stimulation (rTMS) in the acute treatment of poor responder depressed patients: a comparison study between high and low frequency stimulation. *European Psychiatry: The Journal of the Association of European Psychiatrists, 30*(2), 271–276. https://doi.org/https://dx.doi.org/10.1016/j.eurpsy.2014.12.001

Dong, B., Chen, Z., Yin, X., Li, D., Ma, J., Yin, P., … Xu, S. (2017). The efficacy of acupuncture for treating depression-related insomnia compared with a control group: A systematic review and meta-analysis. *BioMed Research International*, https://doi.org/http://dx.doi.org/10.1155/2017/9614810

Duan, D. M., Tu, Y., Chen, L. P., & Wu, Z. J. (2009). Efficacy evaluation for depression with somatic symptoms treated by electroacupuncture combined with Fluoxetine. *Journal of Traditional Chinese Medicine, 29*(3), 167–173.

Duan, D. M., Tu, Y., Jiao, S., & Qin, W. (2011). The relevance between symptoms and magnetic resonance imaging analysis of the hippocampus of depressed patients given electro-acupuncture combined with Fluoxetine intervention - A randomized, controlled trial. *Chinese Journal of Integrative Medicine, 17*(3), 190–199. https://doi.org/https://dx.doi.org/10.1007/s11655-011-0666-6

Dunphy, K., Baker, F. A., Dumaresq, E., Carroll-Haskins, K., Eickholt, J., Ercole, M., … Wosch, T. (2018). Creative arts interventions to address depression in older adults: A systematic review of outcomes, processes, and mechanisms. *Frontiers in Psychology, 9*, 2655. https://doi.org/https://dx.doi.org/10.3389/fpsyg.2018.02655

Falsafi, N. (2016). A Randomized Controlled Trial of Mindfulness Versus Yoga: Effects on Depression and/or Anxiety in College Students. *Journal of the American Psychiatric Nurses Association, 22*(6), 483–497.

Fan, L., Gong, J., Fu, W., Chen, Z., Xu, N., Liu, J., … Xie, H. (2015). Gender-related differences in outcomes on acupuncture and moxibustion treatment among depression patients. *Journal of Alternative & Complementary Medicine, 21*(11), 673–680. https://doi.org/https://dx.doi.org/10.1089/acm.2015.0068

Fava, M., Freeman, M. P., Flynn, M., Hoeppner, B. B., Shelton, R., Iosifescu, D. V, … Papakostas, G. I. (2018). Double-blind, proof-of-concept (POC) trial of Low-Field Magnetic Stimulation (LFMS) augmentation of antidepressant therapy in treatment-resistant depression (TRD). *Brain Stimulation, 11*(1), 75–84. https://doi.org/https://dx.doi.org/10.1016/j.brs.2017.09.010

Feng, Y., Wang, X., Li, S., Zhang, Y., Wang, H., Li, M., …Zhang, Z. (2011). Clinical research of acupuncture on malignant tumor patients for improving depression and sleep quality. *Journal of Traditional Chinese Medicine, 31*(3), 199-202.

Fitzgerald, P. B., Hoy, K. E., Anderson, R. J., & Daskalakis, Z. J. (2016). A study of the pattern of response to rTMS treatment in depression. *Depression & Anxiety, 33*(8), 746–753. https://doi.org/https://dx.doi.org/10.1002/da.22503

Fitzgerald, P. B., Hoy, K., Daskalakis, Z. J., & Kulkarni, J. (2009). A randomized trial of the anti-depressant effects of low- and high-frequency transcranial magnetic stimulation in treatment-resistant depression. *Depression & Anxiety, 26*(3), 229–234. https://doi.org/https://dx.doi.org/10.1002/da.20454

Fitzgerald, P. B., Hoy, K. E., Elliot, D., McQueen, R., Wambeek, L. E., & Daskalakis, Z. J. (2018). Accelerated repetitive transcranial magnetic stimulation in the treatment of depression. *Neuropsychopharmacology, 43*(7), 1565–1572. https://doi.org/http://dx.doi.org/10.1038/s41386-018-0009-9

Fitzgerald, P., Hoy, K., Gunewardene, R., Slack, C., Ibrahim, S., Bailey, M., & Daskalakis, Z. (2011). A randomized trial of unilateral and bilateral prefrontal cortex transcranial magnetic stimulation in treatment-resistant major depression. *Psychological Medicine, 41*(6), 1187–1196. https://doi.org/http://dx.doi.org/10.1017/S0033291710001923

Fitzgerald, P. B., Hoy, K. E., Herring, S. E., McQueen, S., Peachey, A. V, Segrave, R. A., … Daskalakis, Z. (2012). A double blind randomized trial of unilateral left and bilateral prefrontal cortex transcranial magnetic stimulation in treatment resistant major depression. *Journal of Affective Disorders, 139*(2), 193–198. https://doi.org/http://dx.doi.org/10.1016/j.jad.2012.02.017

Fitzgerald, P. B., Hoy, K. E., Singh, A., Gunewardene, R., Slack, C., Ibrahim, S., … Daskalakis, Z. (2013). Equivalent beneficial effects of unilateral and bilateral prefrontal cortex transcranial magnetic stimulation in a large randomized trial in treatment-resistant major depression. *International Journal of Neuropsychopharmacology, 16*(9), 1975–1984. https://doi.org/http://dx.doi.org/10.1017/S1461145713000369

Galletly, C., Gill, S., Clarke, P., Burton, C., & Fitzgerald, P. (2012). A randomized trial comparing repetitive transcranial magnetic stimulation given 3 days/week and 5 days/week for the treatment of major depression: Is efficacy related to the duration of treatment or the number of treatments? *Psychological Medicine, 42*(5), 981–988. https://doi.org/http://dx.doi.org/10.1017/S0033291711001760

Gangadhar, B. N., Naveen, G. H., Rao, M. G., Thirthalli, J., & Varambally, S. (2013). Positive antidepressant effects of generic yoga in depressive out-patients: a comparative study. *Indian Journal of Psychiatry, 55*(7), S369–S373.

Gartlehner, G., Gaynes, B. N., Amick, H. R., Asher, G., Morgan, L. C., Coker-Schwimmer, E., … Lohr, K. N. (2016). Comparative benefits and harms of antidepressant, psychological, complementary, and exercise treatments for major depression: An evidence report for a clinical practice guideline from the American College of Physicians. *Agency for Healthcare Research and Quality, 15*(16), 12.

Gartlehner, G., Gaynes, B. N., Amick, H. R., Asher, G., Morgan, L. C., Coker-Schwimmer, E., … Lohr, K. N. (2015). Nonpharmacological versus pharmacological treatments for adult patients with major depressive disorder. *Agency for Healthcare Research and Quality, 15*(16), 12.

Gartlehner, G., Wagner, G., Matyas, N., Titscher, V., Greimel, J., Lux, L., …Lohr, K. (2017). Pharmacological and non-pharmacological treatments for major depressive disorder: Review of systematic reviews. *BMJ Open, 7*, e014912.

Gaynes, B. N., Lloyd, S. W., Lux, L., Gartlehner, G., Hansen, R. A., Brode, S., … Lohr, K. N. (2014). Repetitive transcranial magnetic stimulation for treatment-resistant depression: A systematic review and meta-analysis. *Journal of Clinical Psychiatry, 75*(5), 477–489. https://doi.org/http://dx.doi.org/10.4088/JCP.13r08815

George, M. S., Lisanby, S. H., Avery, D., McDonald, W. M., Durkalski, V., Pavlicova, M., … Sackeim, H. A. (2010). Daily left prefrontal transcranial magnetic stimulation therapy for major depressive disorder: A sham-controlled randomized trial. *Archives of General Psychiatry, 67*(5), 507–516. https://doi.org/https://dx.doi.org/10.1001/archgenpsychiatry.2010.46

Greer, T. L., Trombello, J. M., Rethorst, C. D., Carmody, T. J., Jha, M. K., Liao, A., … Trivedi, M. H. (2016). Improvements in psychosocial functioning and health-related quality of life following exercise augmentation in patients with treatment response but nonremitted major depressive disorder: Results from the tread study. *Depression & Anxiety*, *33*(9), 870–881. https://doi.org/https://dx.doi.org/10.1002/da.22521

Haesebaert, F., Moirand, R., Schott-Pethelaz, A. M., Brunelin, J., & Poulet, E. (2018). Usefulness of repetitive transcranial magnetic stimulation as a maintenance treatment in patients with major depression. *World Journal of Biological Psychiatry, 19*(1), 74–78. https://doi.org/https://dx.doi.org/10.1080/15622975.2016.1255353

Heinzel, S., Lawrence, J. B., Kallies, G., Rapp, M. A., & Heissel, A. (2015). Using exercise to fight depression in older adults: A systematic review and meta-analysis. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry*, *28*(4), 149–162. https://doi.org/http://dx.doi.org/10.1024/1662-9647/a000133

Helgadottir, B., Hallgren, M., Ekblom, O., & Forsell, Y. (2016). Training fast or slow? Exercise for depression: A randomized controlled trial. *Preventive Medicine*, *91*, 123–131. https://doi.org/10.1016/j.ypmed.2016.08.011

Holvast, F., Massoudi, B., Oude Voshaar, R. C., & Verhaak, P. F. M. (2017). Non-pharmacological treatment for depressed older patients in primary care: A systematic review and meta-analysis. *PLoS ONE*, *12*(9 PG-e0184666), e0184666. https://doi.org/https://dx.doi.org/10.1371/journal.pone.0184666

Hopton, A., Macpherson, H., Keding, A., & Morley, S. (2014). Acupuncture, counselling or usual care for depression and comorbid pain: Secondary analysis of a randomised controlled trial. *BMJ Open, 4*(5), e004964. https://doi.org/https://dx.doi.org/10.1136/bmjopen-2014-004964

Hou, W. H., Chiang, P. T., Hsu, T. Y., Chiu, S. Y., & Yen, Y. C. (2010). Treatment effects of massage therapy in depressed people: A meta-analysis. *Journal of Clinical Psychiatry, 71*(7), 894–901. https://doi.org/https://dx.doi.org/10.4088/JCP.09r05009blu

Huang, T. T., Liu, C. B., Tsai, Y. H., Chin, Y. F., & Wong, C. H. (2015). Physical fitness exercise versus cognitive behavior therapy on reducing the depressive symptoms among community-dwelling elderly adults: A randomized controlled trial. *International Journal of Nursing Studies*, *52*(10), 1542–1552. https://doi.org/https://dx.doi.org/10.1016/j.ijnurstu.2015.05.013

Huang, M. L., Luo, B. Y., Hu, J. B., Wang, S. S., Zhou, W. H., Wei, N., … Xu, Y. (2012). Repetitive transcranial magnetic stimulation in combination with citalopram in young patients with first-episode major depressive disorder: a double-blind, randomized, sham-controlled trial. *Australian & New Zealand Journal of Psychiatry, 46*(3), 257–264. https://doi.org/https://dx.doi.org/10.1177/0004867411433216

Ionson, E., Limbachia, J., Rej, S., Puka, K., Newman, R. I., Wetmore, S., … Vasudev, A. (2018). Effects of Sahaj Samadhi meditation on heart rate variability and depressive symptoms in patients with late-life depression. *British Journal of Psychiatry*, 1–7. https://doi.org/https://dx.doi.org/10.1192/bjp.2018.265

Janicak, P. G., Nahas, Z., Lisanby, S. H., Solvason, H., Sampson, S. M., McDonald, W. M., … Schatzberg, A. F. (2010). Durability of clinical benefit with transcranial magnetic stimulation (TMS) in the treatment of pharmacoresistant major depression: Assessment of relapse during a 6-month, multisite, open-label study. *Brain Stimulation, 3*(4), 187–199. https://doi.org/http://dx.doi.org/10.1016/j.brs.2010.07.003

Josefsson, T., Lindwall, M., & Archer, T. (2014). Physical exercise intervention in depressive disorders: Meta-analysis and systematic review. *Scandinavian Journal of Medicine & Science in Sports*, *24*(2), 259–272. https://doi.org/http://dx.doi.org/10.1111/sms.12050

Kaster, T. S., Daskalakis, Z. J., Noda, Y., Knyahnytska, Y., Downar, J., Rajji, T. K., … Blumberger, D. M. (2018). Efficacy, tolerability, and cognitive effects of deep transcranial magnetic stimulation for late-life depression: A prospective randomized controlled trial. *Neuropsychopharmacology, 43*(11), 2231–2238.

Kaur, M., Chun Lee, R. S., Fitzgerald, P. B., Duffy, S. L., Wong, C., Scott, E. M., & Hickie, I. B. (2015). The relationship between treatment parameters and efficacy with repetitive transcranial magnetic stimulation in depression: A meta-analytic study. *Brain Stimulation, 8* (2), 423–424.

Kavanaugh, B. C., Aaronson, S. T., Clarke, G. N., Holtzheimer, P. E., Johnson, C. W., McDonald, W. M., … Carpenter, L. L. (2018). Neurocognitive effects of repetitive Transcranial Magnetic Stimulation with a 2-coil device in treatment-resistant Major Depressive Disorder. *Journal of ECT, 34*(4) 258–265.

Kedzior, K. K., Gellersen, H. M., Roth, Y., & Zangen, A. (2015). Acute reduction in anxiety after deep transcranial magnetic stimulation (DTMS) in unipolar major depression- a systematic review and meta-analysis. *Psychiatry Research, 230*(3), 971–974. https://doi.org/https://dx.doi.org/10.1016/j.psychres.2015.11.032

Kedzior, K. K., Muller, C., & Gellersen, H. M. (2017). Comparing depression severity after high-frequency repetitive transcranial magnetic stimulation (rTMS) with h1-coil vs. Figure-of-eight (F8)-coil in unipolar major depression: A systematic review and meta-analysis. *Brain Stimulation, 10* (2), 414. https://doi.org/http://dx.doi.org/10.1016/j.brs.2017.01.229

Kedzior, K. K., Muller, C., Gerkensmeier, I., Gellersen, H. M., & Schuchinsky, M. (2017). Predictors of clinical response to deep transcranial magnetic stimulation (dTMS) in unipolar major depression: A systematic review and meta-analysis. *Brain Stimulation, 10*(2), 354–355. https://doi.org/http://dx.doi.org/10.1016/j.brs.2017.01.042

Kerling, A., Kuck, M., Tegtbur, U., Grams, L., Weber-Spickschen, S., Hanke, A., … Kahl, K. (2017). Exercise increases serum brain-derived neurotrophic factor in patients with major depressive disorder. *Journal of Affective Disorders*, *215*, 152–155. https://doi.org/http://dx.doi.org/10.1016/j.jad.2017.03.034

Kerse, N., Hayman, K. J., Moyes, S. A., Peri, K., Robinson, E., Dowell, A., … Arroll, B. (2010). Home-based activity program for older people with depressive symptoms: DeLLITE--A randomized controlled trial. *Annals of Family Medicine*, *8*(3), 214–223. https://doi.org/https://dx.doi.org/10.1370/afm.1093

Kim, D. R., Wang, E., McGeehan, B., Snell, J., Ewing, G., Iannelli, C., … Epperson, C. (2019). Randomized controlled trial of transcranial magnetic stimulation in pregnant women with major depressive disorder. *Brain Stimulation, 12*(1), 96–102. https://doi.org/http://dx.doi.org/10.1016/j.brs.2018.09.005

Kinser, P., Bourguignon, C., Whaley, D., Hauenstein, E., & Taylor, A. (2013). Feasibility, acceptability, and effects of gentle Hatha yoga for women with major depression: Findings from a randomized controlled mixed-methods study. *Archives of Psychiatric Nursing, 27*(3), 137-147.

Kinser, P. A., Elswick, R. K., & Kornstein, S. (2014). Potential long-term effects of a mind-body intervention for women with major depressive disorder: sustained mental health improvements with a pilot yoga intervention. *Archives of Psychiatric Nursing, 28*(6), 377–383. https://doi.org/https://dx.doi.org/10.1016/j.apnu.2014.08.014

Klainin-Yobas, P., Oo, W. N., Suzanne Yew, P. Y., & Lau, Y. (2015). Effects of relaxation interventions on depression and anxiety among older adults: A systematic review. *Aging & Mental Health, 19*(12), 1043–1055. https://doi.org/http://dx.doi.org/10.1080/13607863.2014.997191

Kreuzer, P. M., Schecklmann, M., Lehner, A., Wetter, T. C., Poeppl, T. B., Rupprecht, R., … Langguth, B. (2015). The ACDC pilot trial: targeting the anterior cingulate by double cone coil rTMS for the

treatment of depression. *Brain Stimulation, 8*(2), 240–246. https://doi.org/https://dx.doi.org/10.1016/j.brs.2014.11.014

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

Kumar Dr, G., & Singh, B. K. (2013). To access the role of music therapy in depression and their comparison with drug therapy. *International Journal of Pharmaceutical Sciences and Research, 4*(8), 3099–3102.

Leggett, L. E., Soril, L. J. J., Coward, S., Lorenzetti, D. L., Mackean, G., & Clement, F. M. (2015). Repetitive transcranial magnetic stimulation for treatment-resistant depression in adult and youth populations: A systematic literature review and meta-analysis. *Primary Care Companion to the Journal of Clinical Psychiatry, 17*(6), 379–388. https://doi.org/http://dx.doi.org/10.4088/PCC.15r01807

Leubner, D., & Hinterberger, T. (2017). Reviewing the effectiveness of music interventions in treating depression. *Frontiers in Psychology, 8*, 1109. https://doi.org/https://dx.doi.org/10.3389/fpsyg.2017.01109

Leuchter, A., Cook, I., Feifel, D., Goethe, J., Husain, M., Carpenter, L., … George, M. (2014). Efficacy and safety of low-field synchronized transcranial magnetic stimulation (sTMS) for treatment of major depression. *Neuropsychopharmacology, 39*(7)

Levkovitz, Y., Harel, E. V, Roth, Y., Braw, Y., Most, D., Katz, L. N., … Zangen, A. (2009). Deep transcranial magnetic stimulation over the prefrontal cortex: evaluation of antidepressant and cognitive effects in depressive patients. *Brain Stimulation, 2*(4), 188–200. https://doi.org/https://dx.doi.org/10.1016/j.brs.2009.08.002

Levkovitz, Y., Isserles, M., Padberg, F., Lisanby, S. H., Bystritsky, A., Xia, G., … Zangen, A. (2015). Efficacy and safety of deep transcranial magnetic stimulation for major depression: a prospective multicenter randomized controlled trial. *World Psychiatry, 14*(1), 64–73. https://doi.org/https://dx.doi.org/10.1002/wps.20199

Levkovitz, Y., Sheer, A., Harel, E. V, Katz, L. N., Most, D., Zangen, A., & Isserles, M. (2011). Differential effects of deep TMS of the prefrontal cortex on apathy and depression. *Brain Stimulation, 4*(4), 266–274. https://doi.org/https://dx.doi.org/10.1016/j.brs.2010.12.004

Li, S., Li, Z. F., Wu, Q., Guo, X. C., Xu, Z. H., Li, X. B., … Fu, W. B. (2018). A multicenter, randomized, controlled trial of electroacupuncture for perimenopause women with mild-moderate depression. *BioMed Research International,* https://doi.org/https://dx.doi.org/10.1155/2018/5351210

Liao, S. J., Tan, M. P., Chong, M. C., & Chua, Y. P. (2018). The impact of combined music and tai chi on depressive symptoms among community-dwelling older persons: A cluster randomized controlled

trial. *Issues in Mental Health Nursing, 39*(5), 398–402. https://doi.org/https://dx.doi.org/10.1080/01612840.2017.1417519

Lisanby, S. H., Husain, M. M., Rosenquist, P. B., Maixner, D., Gutierrez, R., Krystal, A., … George, M. S. (2009). Daily left prefrontal repetitive transcranial magnetic stimulation in the acute treatment of major depression: clinical predictors of outcome in a multisite, randomized controlled clinical trial. *Neuropsychopharmacology, 34*(2), 522–534. https://doi.org/https://dx.doi.org/10.1038/npp.2008.118

Liu, X., Clark, J., Siskind, D., Williams, G. M., Byrne, G., Yang, J. L., & Doi, S. A. (2015). A systematic review and meta-analysis of the effects of Qigong and Tai Chi for depressive symptoms. *Complementary Therapies in Medicine, 23*(4), 516–534. https://doi.org/https://dx.doi.org/10.1016/j.ctim.2015.05.001

Liu, Y., Feng, H., Mo, Y., Gao, J., Mao, H., Song, M., … Liu, W. (2015). Effect of soothing-liver and nourishing-heart acupuncture on early selective serotonin reuptake inhibitor treatment onset for depressive disorder and related indicators of neuroimmunology: A randomized controlled clinical trial. *Journal of Traditional Chinese Medicine, 35*(5), 507–513

Lo, H. H., Ng, S. M., Chan, C. L., Lam, K. F., & Lau, B. H. (2013). The Chinese medicine construct "stagnation" in mind-body connection mediates the effects of mindfulness training on depression and anxiety. *Complementary Therapies in Medicine, 21*(4), 348–357. https://doi.org/https://dx.doi.org/10.1016/j.ctim.2013.05.008

Luttenberger, K., Stelzer, E.-M., Forst, S., Schopper, M., Kornhuber, J., & Book, S. (2015). Indoor rock climbing (bouldering) as a new treatment for depression: Study design of a waitlist-controlled randomized group pilot study and the first results. *BMC Psychiatry Vol 15*

Ma, S., Qu, S., Huang, Y., Chen, J., Lin, R., Wang, C., …Zhang, Z. (2012). Improvement in quality of life in depressed patients following verum acupuncture or electroacupuncture plus paroxetine: A randomized controlled study of 157 cases. *Neural Regeneration Research, 7*(27), 2123-2129.

MacPherson, H., Richmond, S., Bland, M., Brealey, S., Gabe, R., Hopton, A., … Watt, I. (2013). Acupuncture and counselling for depression in primary care: A randomised controlled trial. *PLoS, 10*(9), e1001518. https://doi.org/https://dx.doi.org/10.1371/journal.pmed.1001518

Manber, R., Schnyer, R. N., Lyell, D., Chambers, A. S., Caughey, A. B., Druzin, M., … Allen, J. J. (2010). Acupuncture for depression during pregnancy: A randomized controlled trial. *Obstetrics & Gynecology, 115*(3), 511–520. https://doi.org/https://dx.doi.org/10.1097/AOG.0b013e3181cc0816

Maneeton, B., Maneeton, N., Woottiluk, P., & Oonarom, A. (2019). Repetitive transcranial magnetic stimulation combined with antidepressants for major depressive disorder: A meta-analysis and systematic review. *European Neuropsychopharmacology, 29* (1), S65–S66. https://doi.org/http://dx.doi.org/10.1016/j.euroneuro.2018.11.1043

Mantovani, A., Aly, M., Dagan, Y., Allart, A., & Lisanby, S. H. (2013). Randomized sham controlled trial of repetitive transcranial magnetic stimulation to the dorsolateral prefrontal cortex for the treatment of panic disorder with comorbid major depression. *Journal of Affective Disorders, 144*(1-2), 153–159. https://doi.org/https://dx.doi.org/10.1016/j.jad.2012.05.038

Martinot, M.-L. P., Galinowski, A., Ringuenet, D., Gallarda, T., Lefaucheur, J.-P., Bellivier, F., … Martinot, J.-L. (2010). Influence of prefrontal target region on the efficacy of repetitive transcranial magnetic stimulation in patients with medication-resistant depression: A. *International Journal of Neuropsychopharmacology, 13*(1), 45–59. https://doi.org/http://dx.doi.org/10.1017/S146114570900008X

Martiny, K., Refsgaard, E., Lund, V., Lunde, M., Thougaard, B., Lindberg, L., & Bech, P. (2015). Maintained superiority of chronotherapeutics vs. exercise in a 20-week randomized follow-up trial in major depression. *Acta Psychiatrica Scandinavica*, *131*(6), 446–457. https://doi.org/https://dx.doi.org/10.1111/acps.12402

Mead, G. E., Morley, W., Campbell, P., Greig, C. A., McMurdo, M., & Lawlor, D. A. (2009). Exercise for depression. *Cochrane Database of Systematic Reviews*, (3), CD004366. https://doi.org/https://dx.doi.org/10.1002/14651858.CD004366.pub4

Meyer, J. D., Koltyn, K. F., Stegner, A. J., Kim, J. S., & Cook, D. B. (2016). Influence of exercise intensity for improving depressed mood in depression: A dose-response study. *Behavior Therapy*, *47*(4), 527–537. https://doi.org/https://dx.doi.org/10.1016/j.beth.2016.04.003

Meyer, J. D., Koltyn, K. F., Stegner, A. J., Kim, J. S., & Cook, D. B. (2016). Relationships between serum BDNF and the antidepressant effect of acute exercise in depressed women. *Psychoneuroendocrinology*, *74*, 286–294. https://doi.org/https://dx.doi.org/10.1016/j.psyneuen.2016.09.022

Minghetti, A., Faude, O., Hanssen, H., Zahner, L., Gerber, M., & Donath, L. (2018). Sprint interval training (SIT) substantially reduces depressive symptoms in major depressive disorder (MDD): A randomized controlled trial. *Psychiatry Research*, *265*, 292–297.

Mitchell, J., Field, T., Diego, M., Bendell, D., Newton, R., & Pelaez, M. (2012). Yoga reduces prenatal depression symptoms. *Psychology, 13*, 782-786.

Mogg, A., Pluck, G., Eranti, S., Landau, S., Purvis, R., Brown, R., … McLoughlin, D. (2008). A randomized controlled trial with 4-month follow-up of adjunctive repetitive transcranial magnetic stimulation of the left prefrontal cortex for depression. *Psychological Medicine, 38*(3), 323–333. https://doi.org/http://dx.doi.org/10.1017/S0033291707001663

Morres, I. D., Hatzigeorgiadis, A., Stathi, A., Comoutos, N., Arpin-Cribbie, C., Krommidas, C., & Theodorakis, Y. (2019). Aerobic exercise for adult patients with major depressive disorder in mental health services: A systematic review and meta-analysis. *Depression & Anxiety*, *36*(1), 39–53. https://doi.org/https://dx.doi.org/10.1002/da.22842

Mura, G., & Carta, M. G. (2013). Physical activity in depressed elderly. A systematic review. *Clinical Practice & Epidemiology in Mental Health*, *9*, 125–135. https://doi.org/https://dx.doi.org/10.2174/1745017901309010125

Mutz, J., Edgcumbe, D., Thoma, V., & Fu, C. H. Y. (2017). Non-invasive brain stimulation for the treatment of depression: Systematic review and meta-analysis of randomised, sham-controlled trials. *European Neuropsychopharmacology, 27*(4), S811.

Nan, J. K., & Ho, R. T. (2017). Effects of clay art therapy on adults outpatients with major depressive disorder: A randomized controlled trial. *Journal of Affective Disorders, 217*, 237–245. https://doi.org/http://dx.doi.org/10.1016/j.jad.2017.04.013

Naveen, G., Varambally, S., Thirthalli, J., Rao, M., Christopher, R., & Gangadhar, B. (2016). Serum cortisol and BDNF in patients with major depression-Effect of yoga. *International Review of Psychiatry, 28*(3), 273–278. https://doi.org/http://dx.doi.org/10.1080/09540261.2016.1175419

Neunhauserer, D., Sturm, J., Baumgartlinger, M. M., Niederseer, D., Ledl-Kurkowski, E., Steidle, E., … Niebauer, J. (2013). Hiking in suicidal patients: neutral effects on markers of suicidality. *American Journal of Medicine*, *126*(10), 927–930. https://doi.org/https://dx.doi.org/10.1016/j.amjmed.2013.05.008

Neviani, F., Belvederi Murri, M., Mussi, C., Triolo, F., Toni, G., Simoncini, E., … Neri, M. (2017). Physical exercise for late life depression: effects on cognition and disability. *International Psychogeriatrics*, *29*(7), 1105–1112. https://doi.org/https://dx.doi.org/10.1017/S1041610217000576

Nguyen, S., Marshak, H., Dos Santos, H., Luu, S., Berk, L., McMahon, P., & Riggs, M. (2014). The effects of simultaneous exercise and psychotherapy on depressive symptoms in inpatient, psychiatric older adults. *Advances, 28*(4), 8-17.

Nongpiur, A., Sinha, V. K., Praharaj, S. K., & Goyal, N. (2011). Theta-patterned, frequency-modulated priming stimulation enhances low-frequency, right prefrontal cortex repetitive transcranial magnetic stimulation (rTMS) in depression: a randomized, sham-controlled study. *Journal of Neuropsychiatry & Clinical Neurosciences, 23*(3), 348–357. https://doi.org/https://dx.doi.org/10.1176/appi.neuropsych.23.3.348

Pallanti, S., Bernardi, S., Di Rollo, A., Antonini, S., & Quercioli, L. (2010). Unilateral low frequency versus sequential bilateral repetitive transcranial magnetic stimulation: is simpler better for treatment of resistant depression? *Neuroscience, 167*(2), 323–328. https://doi.org/https://dx.doi.org/10.1016/j.neuroscience.2010.01.063

Philip, N. S., Dunner, D. L., Dowd, S. M., Aaronson, S. T., Brock, D. G., Carpenter, L. L., … George, M. S. (2016). Can medication free, treatment-resistant, depressed patients who initially respond to TMS be maintained off medications? A prospective, 12-month multisite randomized pilot study. *Brain Stimulation, 9*(2), 251–257. https://doi.org/https://dx.doi.org/10.1016/j.brs.2015.11.007

Plewnia, C., Pasqualetti, P., Groe, S., Schlipf, S., Wasserka, B., Zwissler, B., & Fallgatter, A. (2014). Treatment of major depression with bilateral theta burst stimulation: a randomized controlled pilot trial. *Journal of Affective Disorders, 156*, 219–223.

Prasser, J., Schecklmann, M., Poeppl, T. B., Frank, E., Kreuzer, P. M., Hajak, G., … Langguth, B. (2015). Bilateral prefrontal rTMS and theta burst TMS as an add-on treatment for depression: a randomized placebo controlled trial. *World Journal of Biological Psychiatry, 16*(1), 57–65. https://doi.org/https://dx.doi.org/10.3109/15622975.2014.964768

Prakhinkit, S., Suppapitiporn, S., Tanaka, H., & Suksom, D. (2014). Effects of Buddhism walking meditation on depression, functional fitness, and endothelium-dependent vasodilation in depressed elderly. *Journal of Alternative & Complementary Medicine, 20*(5), 411–416. https://doi.org/https://dx.doi.org/10.1089/acm.2013.0205

Prathikanti, S., Rivera, R., Cochran, A., Tungol, J. G., Fayazmanesh, N., & Weinmann, E. (2017). Treating major depression with yoga: A prospective, randomized, controlled pilot trial. *PLoS ONE, 12*(3), e0173869. https://doi.org/https://dx.doi.org/10.1371/journal.pone.0173869

Qu, S. S., Huang, Y., Zhang, Z. J., Chen, J. Q., Lin, R. Y., Wang, C. Q., … Zhang, Y. C. (2013). A 6-week randomized controlled trial with 4-week follow-up of acupuncture combined with paroxetine in patients with major depressive disorder. *Journal of Psychiatric Research, 47*(6), 726–732. https://doi.org/https://dx.doi.org/10.1016/j.jpsychires.2013.02.004

Rahman, M. S., Helgadottir, B., Hallgren, M., Forsell, Y., Stubbs, B., Vancampfort, D., & Ekblom, O. (2018). Cardiorespiratory fitness and response to exercise treatment in depression. *BJPsych Open*, *4*(5), 346–351. https://doi.org/https://dx.doi.org/10.1192/bjo.2018.45

Rajagopal, R., Jois, S. N., Mallikarjuna Majgi, S., Anil Kumar, M., & Shashidhar, H. (2018). Amelioration of mild and moderate depression through Pranic Healing as adjuvant therapy: Randomised double-blind controlled trial. *Australasian Psychiatry, 26*(1), 82–87. https://doi.org/http://dx.doi.org/10.1177/1039856217726449

Ray, S., Nizamie, S., Akhtar, S., Praharaj, S. K., Mishra, B. R., & Zia-ul-Haq, M. (2011). Efficacy of adjunctive high frequency repetitive transcranial magnetic stimulation of left prefrontal cortex in depression: A randomized sham controlled study. *Journal of Affective Disorders, 128*(1-2), 153–159. https://doi.org/http://dx.doi.org/10.1016/j.jad.2010.06.027

Razza, L. B., Moffa, A. H., Moreno, M. L., Carvalho, A. F., Padberg, F., Fregni, F., & Brunoni, A. R. (2018). A systematic review and meta-analysis on placebo response to repetitive transcranial magnetic stimulation for depression trials. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 81*, 105–113. https://doi.org/http://dx.doi.org/10.1016/j.pnpbp.2017.10.016

Rethorst, C. D., Landers, D. M., Nagoshi, C. T., & Ross, J. T. (2010). Efficacy of exercise in reducing depressive symptoms across 5-HTTLPR genotypes. *Medicine & Science in Sports & Exercise*, *42*(11), 2141–2147. https://doi.org/https://dx.doi.org/10.1249/MSS.0b013e3181de7d51

Rethorst, C., Sunderajan, P., Greer, T., Grannemann, B., Nakonezny, P., Carmody, T., & Trivedi, M. (2013). Does exercise improve self-reported sleep quality in non-remitted major depressive disorder? *Psychological Medicine*, *43*(4), 699–709. https://doi.org/http://dx.doi.org/10.1017/S0033291712001675

Rethorst, C. D., South, C. C., Rush, A., Greer, T. L., & Trivedi, M. H. (2017). Prediction of treatment outcomes to exercise in patients with nonremitted major depressive disorder. *Depression and Anxiety, 34*(12), 1116–1122. https://doi.org/http://dx.doi.org/10.1002/da.22670

Rhyner, K. T., & Watts, A. (2016). Exercise and depressive symptoms in older adults: A systematic meta-analytic review. *Journal of Aging & Physical Activity, 24*(2), 234–246. https://doi.org/https://dx.doi.org/10.1123/japa.2015-0146

Rimer, J., Dwan, K., Lawlor, D. A., Greig, C. A., McMurdo, M., Morley, W., & Mead, G. E. (2012). Exercise for depression. *Cochrane Database of Systematic Reviews*, (7), CD004366. https://doi.org/https://dx.doi.org/10.1002/14651858.CD004366.pub5

Rossini, D., Lucca, A., Magri, L., Malaguti, A., Smeraldi, E., Colombo, C., & Zanardi, R. (2010). A symptom-specific analysis of the effect of high-frequency left or low-frequency right transcranial magnetic stimulation over the dorsolateral prefrontal cortex in major depression. *Neuropsychobiology, 62*(2), 91–97. https://doi.org/https://dx.doi.org/10.1159/000315439

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

Schutter, D. J., Laman, D. M., van Honk, J., Vergouwen, A. C., & Koerselman, G. F. (2009). Partial clinical response to 2 weeks of 2 Hz repetitive transcranial magnetic stimulation to the right parietal cortex in depression. *International Journal of Neuropsychopharmacology, 12*(5), 643–650. https://doi.org/https://dx.doi.org/10.1017/S1461145708009553

Schuver, K. J., & Lewis, B. A. (2016). Mindfulness-based yoga intervention for women with depression. *Complementary Therapies in Medicine, 26*, 85–91. https://doi.org/https://dx.doi.org/10.1016/j.ctim.2016.03.003

Seo, J. Y., & Chao, Y. Y. (2018). Effects of exercise interventions on depressive symptoms among community-dwelling older adults in the United States: A systematic review. *Journal of Gerontological Nursing*, *44*(3), 31–38. https://doi.org/https://dx.doi.org/10.3928/00989134-20171024-01

Serafini, G., Pompili, M., Belvederi Murri, M., Respino, M., Ghio, L., Girardi, P., … Amore, M. (2015). The effects of repetitive transcranial magnetic stimulation on cognitive performance in treatment-resistant depression. A systematic review. *Neuropsychobiology, 71*(3), 125–139. https://doi.org/http://dx.doi.org/10.1159/000381351

Shahidi, M., Mojtahed, A., Modabbernia, A., Mojtahed, M., Shafiabady, A., Delavar, A., & Honari, H. (2011). Laughter yoga versus group exercise program in elderly depressed women: a randomized controlled trial. *International Journal of Geriatric Psychiatry, 26*(3), 322–327. https://doi.org/https://dx.doi.org/10.1002/gps.2545

Sherwood, A., Blumenthal, J. A., Smith, P. J., Watkins, L. L., Hoffman, B. M., & Hinderliter, A. L. (2016). Effects of exercise and sertraline on measures of coronary heart disease risk in patients with major depression: results from the smile-ii randomized clinical trial. *Psychosomatic Medicine*, *78*(5), 602–609. https://doi.org/https://dx.doi.org/10.1097/PSY.0000000000000301

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

Smith, C. A., Hay, P. P., & Macpherson, H. (2010). Acupuncture for depression. *Cochrane Database of Systematic Reviews*, (1), CD004046. https://doi.org/https://dx.doi.org/10.1002/14651858.CD004046.pub3

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

Streeter, C. C., Gerbarg, P. L., Whitfield, T. H., Owen, L., Johnston, J., Silveri, M. M., … Jensen, J. E. (2017). Treatment of major depressive disorder with Iyengar yoga and coherent breathing: A randomized controlled dosing study. *The Journal of Alternative and Complementary Medicine, 23*(3), 201–207. https://doi.org/http://dx.doi.org/10.1089/acm.2016.0140

Strom, M., Uckelstam, C. J., Andersson, G., Hassmen, P., Umefjord, G., & Carlbring, P. (2013). Internet-delivered therapist-guided physical activity for mild to moderate depression: A randomized controlled trial. *PeerJ*, *1*

Stubbs, B., Vancampfort, D., Rosenbaum, S., Ward, P. B., Richards, J., Ussher, M., & Schuch, F. B. (2016). 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 Medicine*, *46*(5), 699–713. https://doi.org/https://dx.doi.org/10.1007/s40279-015-0441-5

Sun, M., Lanctot, K., Herrmann, N., & Gallagher, D. (2018). Exercise for cognitive symptoms in depression: A systematic review of interventional studies. *Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie*, *63*(2), 115–128. https://doi.org/https://dx.doi.org/10.1177/0706743717738493

Sun, H., Zhao, H., Ma, C., Bao, F., Zhang, J., Wang, D. H., … He, W. (2013). Effects of electroacupuncture on depression and the production of glial cell line-derived neurotrophic factor compared with fluoxetine: a randomized controlled pilot study. *Journal of Alternative & Complementary Medicine, 19*(9), 733–739. https://doi.org/https://dx.doi.org/10.1089/acm.2011.0637

Theleritis, C., Sakkas, P., Paparrigopoulos, T., Vitoratou, S., Tzavara, C., Bonaccorso, S., … Psarros, C. (2017). Two versus one high-frequency repetitive Transcranial Magnetic Stimulation session per day for treatment-resistant depression: A randomized sham-controlled trial. *Journal of ECT, 33*(3), 190–197. https://doi.org/https://dx.doi.org/10.1097/YCT.0000000000000387

Tolahunase, M. R., Sagar, R., Faiq, M., & Dada, R. (2018). Yoga- and meditation-based lifestyle intervention increases neuroplasticity and reduces severity of major depressive disorder: A randomized controlled trial. *Restorative Neurology & Neuroscience, 36*(3), 423–442. https://doi.org/https://dx.doi.org/10.3233/RNN-170810

Toups, M., Carmody, T., Greer, T., Rethorst, C., Grannemann, B., & Trivedi, M. H. (2017). Exercise is an effective treatment for positive valence symptoms in major depression. *Journal of Affective Disorders*, *209*, 188–194. https://doi.org/https://dx.doi.org/10.1016/j.jad.2016.08.058

Triggs, W. J., Ricciuti, N., Ward, H. E., Cheng, J., Bowers, D., Goodman, W. K., … Nadeau, S. E. (2010). Right and left dorsolateral pre-frontal rTMS treatment of refractory depression: A randomized, sham-controlled trial. *Psychiatry Research, 178*(3), 467–474. https://doi.org/https://dx.doi.org/10.1016/j.psychres.2010.05.009

Trimmer, C., Tyo, R., Pikard, J., McKenna, C., & Naeem, F. (2018). Low-intensity cognitive behavioural therapy-based music group (CBT-Music) for the treatment of symptoms of anxiety and depression: A feasibility study. *Behavioural and Cognitive Psychotherapy, 46*(2), 168–181. https://doi.org/http://dx.doi.org/10.1017/S1352465817000480

Trivedi, M. H., Greer, T. L., Church, T. S., Carmody, T. J., Grannemann, B. D., Galper, D. I., … Blair, S. N. (2011). Exercise as an augmentation treatment for nonremitted major depressive disorder: A randomized, parallel dose comparison. *Journal of Clinical Psychiatry, 72*(5), 677–684. https://doi.org/https://dx.doi.org/10.4088/JCP.10m06743

Tsang, H. W., Chan, E. P., & Cheung, W. M. (2008). Effects of mindful and non-mindful exercises on people with depression: A systematic review. *British Journal of Clinical Psychology, 47*(3), 303–322. https://doi.org/https://dx.doi.org/10.1348/014466508X279260

Wang, W. D., Lu, X. Y., Ng, S. M., Hong, L., Zhao, Y., Lin, Y. N., & Wang, F. (2013). Effects of electro-acupuncture on personality traits in depression: a randomized controlled study. *Chinese Journal of Integrative Medicine, 19*(10), 777–782. https://doi.org/https://dx.doi.org/10.1007/s11655-013-1594-4

Wang, T., Wang, L., Tao, W., & Chen, L. (2014). Acupuncture combined with an antidepressant for patients with depression in hospital: A pragmatic randomised controlled trial. *Acupuncture in Medicine, 32*(4), 308–312. https://doi.org/https://dx.doi.org/10.1136/acupmed-2013-010469

Wang, H. N., Wang, X. X., Zhang, R. G., Wang, Y., Cai, M., Zhang, Y. H., … Zhang, Z. J. (2017). Clustered repetitive transcranial magnetic stimulation for the prevention of depressive

relapse/recurrence: a randomized controlled trial. *Translational Psychiatry, 7*(12), 1292. https://doi.org/https://dx.doi.org/10.1038/s41398-017-0001-x

Wen, X., Wu, Q., Liu, J., Xu, Z., Fan, L., Chen, X., … Fu, W. (2018). Randomized single-blind multicenter trial comparing the effects of standard and augmented acupuncture protocols on sleep quality and depressive symptoms in patients with depression. *Psychology Health & Medicine, 23*(4), 375–390. https://doi.org/https://dx.doi.org/10.1080/13548506.2017.1363399

Whiting, M., Leavey, G., Scammell, A., Au, S., & King, M. (2008). Using acupuncture to treat depression: A feasibility study. *Complementary Therapies in Medicine, 16,* 87-91.

Williams, C. & Tappen, R. (2008). Exercise training for depressed older adults with Alzheimer's Disease. *Aging Mental Health, 12*(1), 72-80.

Xie, M., Jiang, W., & Yang, H. (2015). Efficacy and safety of the Chinese herbal medicine shuganjieyu with and without adjunctive repetitive transcranial magnetic stimulation (rTMS) for geriatric depression: a randomized controlled trial. *Shanghai Jingshen Yixue, 27*(2), 103–110. https://doi.org/https://dx.doi.org/10.11919/j.issn.1002-0829.214151

Yang, T. T., Hsiao, F. H., Wang, K. C., Ng, S. M., Ho, R. T., Chan, C. L., … Chen, Y. T. (2009). The effect of psychotherapy added to pharmacotherapy on cortisol responses in outpatients with major depressive disorder. *Journal of Nervous & Mental Disease, 197*(6), 401–406. https://doi.org/https://dx.doi.org/10.1097/NMD.0b013e3181a61594

Yeh, S. H., Lin, L. W., Chuang, Y. K., Liu, C. L., Tsai, L. J., Tsuei, F. S., … Yang, K. D. (2015). Effects of music aerobic exercise on depression and brain-derived neurotrophic factor levels in community dwelling women. *BioMed Research International*, *2015*, 135893. https://doi.org/https://dx.doi.org/10.1155/2015/135893

Yeung, W. F., Chung, K. F., Tso, K. C., Zhang, S. P., Zhang, Z. J., & Ho, L. M. (2011). Electroacupuncture for residual insomnia associated with major depressive disorder: A randomized controlled trial. *Sleep, 34*(6), 807–815. https://doi.org/https://dx.doi.org/10.5665/SLEEP.1056

Yin, J., & Dishman, R. K. (2014). The effect of Tai Chi and Qigong practice on depression and anxiety symptoms: A systematic review and meta-regression analysis of randomized controlled trials. *Mental Health and Physical Activity, 7*(3), 135–146. https://doi.org/http://dx.doi.org/10.1016/j.mhpa.2014.08.001

Yip, A. G., George, M. S., Tendler, A., Roth, Y., Zangen, A., & Carpenter, L. L. (2017). 61% of unmedicated treatment resistant depression patients who did not respond to acute TMS treatment responded after four weeks of twice weekly deep TMS in the Brainsway pivotal trial. *Brain Stimulation, 10*(4), 847–849. https://doi.org/https://dx.doi.org/10.1016/j.brs.2017.02.013

Zanetidou, S., Belvederi Murri, M., Menchetti, M., Toni, G., Asioli, F., Bagnoli, L., … Bertakis, K. D. (2016). Physical Exercise for Late-Life Depression: customizing an Intervention for Primary Care. *Journal of the American Geriatrics Society*

Zhang, Z. J., Ng, R., Man, S. C., Li, J. T. Y., Wong, W., Wong, H. K., … Sze, S. C. W. (2013). Use of electroacupuncture to accelerate the antidepressant action of selective serotonin reuptake inhibitors: a single-blind, randomised, controlled study. *Hong Kong Medical Journal, 19*(6), S12–S16.

Zhang, Z.-J., Ng, R., Man, S. C., Li, T. Y. J., Wong, W., Tan, Q.-R., … Wong, V. T. (2012). Dense cranial electroacupuncture stimulation for major depressive disorder-A single-blind, randomized, controlled study. *PLoS ONE Vol 7*(1)

Zhang, W. J., Yang, X. B., & Zhong, B. L. (2009). Combination of acupuncture and fluoxetine for depression: a randomized, double-blind, sham-controlled trial. *Journal of Alternative & Complementary Medicine, 15*(8), 837–844. https://doi.org/https://dx.doi.org/10.1089/acm.2008.0607

Zhao, K., Bai, Z. G., Bo, A., & Chi, I. (2016). A systematic review and meta-analysis of music therapy for the older adults with depression. *International Journal of Geriatric Psychiatry, 31*(11), 1188–1198. https://doi.org/https://dx.doi.org/10.1002/gps.4494

Zhao, L., Wu, H., Zhou, X., Wang, Q., Zhu, W., & Chen, J. (2012). Effects of progressive muscular relaxation training on anxiety, depression and quality of life of endometriosis patients under gonadotrophin-releasing hormone agonist therapy. *European Journal of Obstetrics, Gynecology, & Reproductive Biology, 162*(2), 211–215. https://doi.org/https://dx.doi.org/10.1016/j.ejogrb.2012.02.029

Appendix C

See **Figures 4 and 5** below for bubble maps. Bubble maps provide a visual overview of the distribution of evidence for the complementary and integrative health and other interventions included in these systematic reviews. The bubble maps display information about the research meeting the inclusion and exclusion criteria (see Appendix A) for these reviews and include the following:

- **The strength of evidence** (y-axis)
  - The y-axis provides an overview of the quantity of research for an intervention. For this estimate, we used the number of individual RCTs and/or the number of RCTs included in previously published systematic reviews. The color of the bubbles indicates the strength of evidence (SOE). The lighter the color of a bubble, the higher the SOE and vice versa.
- **The direction of findings** (x-axis)
  - The x-axis provides an estimate of the clinical effectiveness of an intervention with the bubble maps differentiating the findings with three different categories, which are, "favors control"; "no difference"; and "favors intervention". Control groups are important to consider and have been noted in the maps as well, given that some studies have an active control and others do not.
- **The confidence in the reported effect** (bubble size)
  - The size of a bubble indicates the level of confidence in the reported effect. Next to each bubble we abbreviate the intervention, the control group, and note the number of studies conducted.

It is important to note that, due to the number of studies included and the scope of these systematic reviews, the bubble maps may only represent limited information.

Figure 4. Bubble Plot of Findings for Depression Symptoms

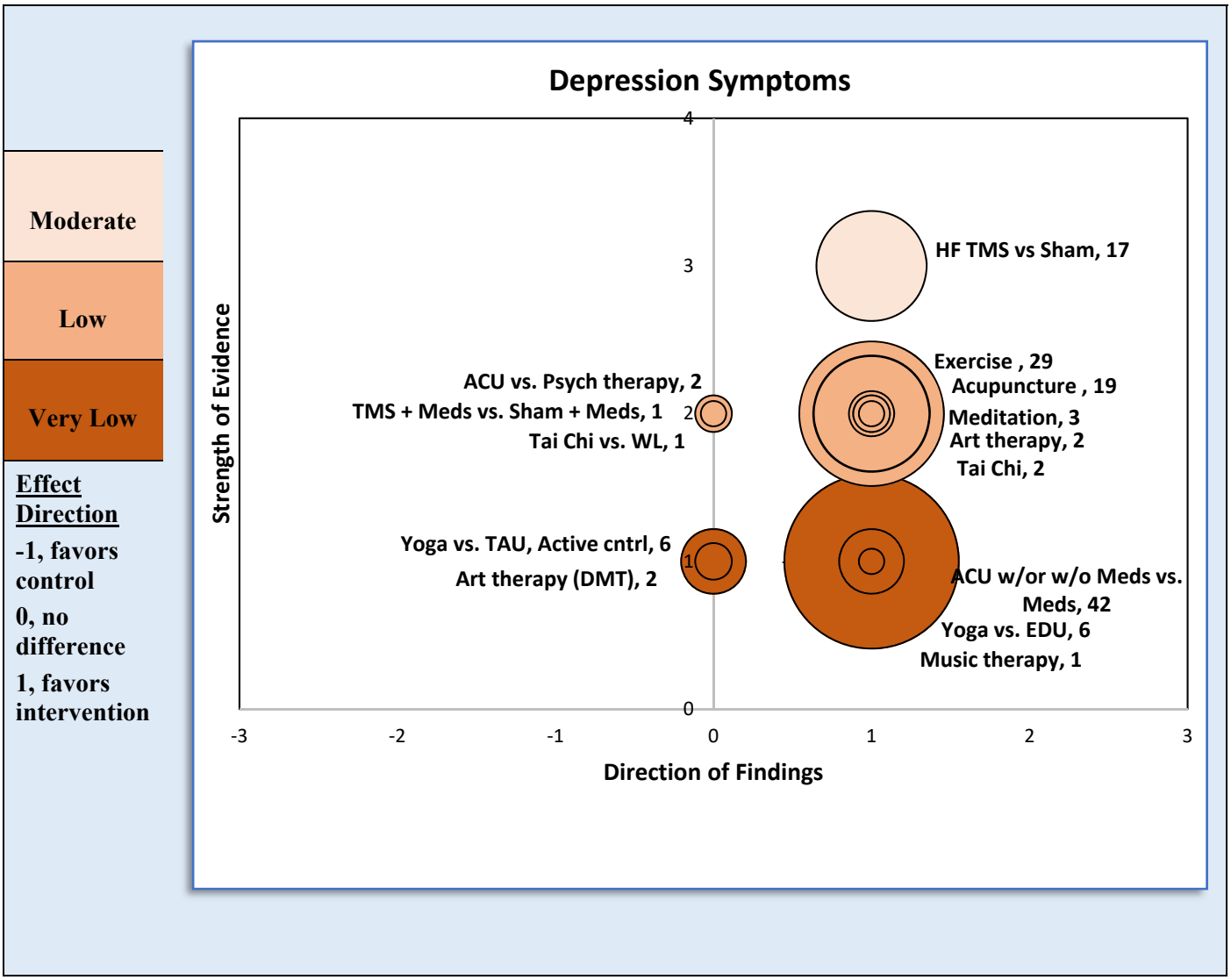


Figure 5. Bubble Plot of Findings for Remission