Aerobic, resistance, and mind-body exercise are equivalent to mitigate symptoms of depression in older adults: A systematic review and network meta-analysis of randomised controlled trials [version 1; peer review: 1 approved, 2 approved with reservations]

Kyle J. Miller¹, Pinyadapat Areerob¹, Declan Hennessy²¹, Daniela C. Gonçalves-Bradley², Christopher Mesagno¹, Fergal Grace¹

¹School of Health and Life Sciences, Federation University, Ballarat, Victoria, 3350, Australia
²Nuffield Department of Population Health, University of Oxford, Oxford, UK

Abstract

Background: Exercise has been identified as an allied health strategy that can support the management of depression in older adults, yet the relative effectiveness for different exercise modalities is unknown. To meet this gap in knowledge, we present a systematic review and network meta-analysis of randomised controlled trials (RCTs) to examine the head-to-head effectiveness of aerobic, resistance, and mind-body exercise to mitigate depressive symptoms in adults aged ≥ 65 years.

Methods: A PRISMA-NMA compliant review was undertaken on RCTs from inception to September 12th, 2019. PubMed, Web of Science, CINAHL, Health Source: Nursing/Academic Edition, PsycARTICLES, PsycINFO, and SPORTDiscus were systematically searched for eligible RCTs enrolling adults with a mean age ≥ 65 years, comparing one or more exercise intervention arms, and which used valid measures of depressive symptomology. Comparative effectiveness was evaluated using network meta-analysis to combine direct and indirect evidence, controlling for inherent variation in trial control groups.

Results: The systematic review included 81 RCTs, with 69 meeting eligibility for the network meta-analysis (n = 5,379 participants). Pooled analysis found each exercise type to be effective compared with controls (Hedges' $g = -0.27$ to $-0.51$). Relative head-to-head comparisons were statistically comparable between exercise types: resistance versus aerobic (Hedges' $g = -0.06$, $PrI = -0.91$, 0.79), mind-body versus aerobic (Hedges' $g = -0.12$, $PrI = -0.95$, 0.72), mind-body
versus resistance (Hedges' $g = -0.06$, $P_{f1} = -0.90$, 0.79). High levels of compliance were demonstrated for each exercise treatment.

**Conclusions:** Aerobic, resistance, and mind-body exercise demonstrate equivalence to mitigate symptoms of depression in older adults aged $\geq 65$ years, with comparably encouraging levels of compliance to exercise treatment. These findings coalesce with previous findings in clinically depressed older adults to encourage personal preference when prescribing exercise for depressive symptoms in older adults, irrespective of severity.

**Registration:** PROSPERO CRD42018115866 (23/11/2018).

**Keywords**
Older adults, elderly, seniors, exercise, physical activity, depression, randomised controlled trial, RCT
Introduction
At the close of this decade, the last remaining ‘baby boomers’ will transition to an expanding peer demographic aged ≥ 65 years projected to constitute more than one billion older adults, worldwide. Physical exercise is proposed as a low-risk adjunctive mitigant of age-associated functional deterioration in mental health, including for dementia and depression. In light of impending demographic shifts, and with the burden of age-associated depression estimated to affect ~20% of older adults, 2030 may confer a burden of 200 million adults aged ≥ 65 years presenting with clinical depression. In preparation for inevitable future demands on primary care systems, there is prevailing opportunity for concerted efforts to support primary and allied health personnel with informed preventative strategies.

International public health consortia are in concert with the antidepressant effects of exercise as a low-risk adjunct for optimal mental health. While we do not yet have the answers for low uptake of exercise in older adults, it may in some way be due to nuanced regimen design, which in turn, may similarly impact compliance in exercise prescription by primary and stakeholders in aged care. By exemplar, ‘Exercise is Medicine’ is a global initiative promulgated through 40 member countries by the American College of Sports Medicine with a platform to promote and encourage routine physical exercise for general health and a broad range of medical conditions. This manifesto encourages primary care physicians and health care providers to refer patients to qualified exercise personnel when prescribing treatment plans. However, despite conventional agreement for exercise as a prophylactic for geriatric depression, contemporary literature is yet to quantify the antidepressant treatment effectiveness for individual exercise types.

Perhaps a lesser appreciated obstacle to optimising exercise prescription for mental health in older adults lies with the ‘catch all’ characterisation of exercise. During the past four decades, widely different metabolic, social, and environmental demands between exercise modalities (i.e., running vs. weightlifting vs. Tai Chi) have been well-characterised. Given that there is variation between exercise regimens, and these variations are not merely semantics, one may be surprised to discover that only a few randomised controlled trials (RCTs) have deliberately compared the antidepressant effects of different exercise regimens in older adults. This begs a meaningful question, ‘are all exercise types equal?’.

In attempting to quantify the magnitude of potential antidepressant effects of exercise, researchers have deployed conventional pairwise meta-analysis, during recent years. In departure from the value offered by pooling conventional treatments during pairwise meta-analysis, this amalgamation is inherently prone to overgeneralisation and concomitant overestimation of treatment effectiveness. More specifically in mental health literature, pooling individual trial effect comparisons during the pairwise meta-analytical process precludes any opportunity for head-to-head comparison between different exercise types. In the same vein, a further nuance of pairwise meta-analysis is that effect-sensitivity is compromised by foregoing potentially relevant characterisation of control arms (i.e., wait-list, usual care, attention-control) of included trials. In this respect, one must acknowledge that pairwise meta-analysis has intrinsic restrictive boundaries in circumstances where head-to-head comparison of different exercise treatments are required.

In circumstances where relative effectiveness of multiple treatment comparisons are required, network meta-analysis offers a methodological solution to provide more precise pooled head-to-head effect estimates than may otherwise be achieved. Performed correctly, network meta-analysis can allow comparative effectiveness of exercise treatment regimens (aerobic, resistance, and mind-body), avoiding the assumption of treatment homogeneity, controlling for bias from small study effects, and likewise, controlling for characteristically different control arms (wait-list, usual care, and attention-control).

Exercise regimens are broadly categorised into either aerobic, resistance, or mind-body exercise types, and any individual exercise program may consist of multiple combinations of these activities. Extended description of these three exercise types may be found as Extended data. It is well established that independent regimens of either aerobic or resistance exercise elicit uniquely different metabolic and phenotypic adaptations, whereas mind-body exercise is unique as a low impact form of exercise. Therefore, it is reasonable to theorise comparative differences in the capacity to elicit any antidepressant effect. Until recently, investigation of head-to-head treatment effects of aerobic, resistance, and mind-body exercise had not been undertaken and their comparative ability to moderate symptoms of depression in older adults without pre-existing clinical depression is unknown.

Recently, Miller and colleagues quantitatively compared head-to-head effectiveness of aerobic, resistance, and mind-body exercise in older adults with pre-existing clinical depression. In realisation of their study hypothesis, Miller et al. excluded studies of older adults which lacked participant-level diagnosis of clinical depression. In doing so, there is an assumption that older adults with diagnosis of clinical depression should not be pooled with similarly aged counterparts without diagnosis. These differences are not merely semantics, and while beyond the broad scope of the present study, are worthy of summary.

Present day phenotypic characterisation of clinically depression in older adults is supported by the coalition of more than six decades of cross-disciplinary research into major depressive disorders. Limbic brain regions, monoamine neurotransmitters, and the hypothalamic-pituitary-adrenal (HPA) axis contribute to the pathophysiology of depression. In addition, hypothalamic oversight of pulsatile stress hormone perception from brain regions is central for homeostatic regulation of human stress response and maintenance of systemic feedback loop physiology in concert with the pituitary and adrenal cortex. Similarly, it is well-established that stress hormones have broad infiltration through large areas of the brain via neuronal supply originating from midbrain and brainstem regions. Indeed, these monoaminergic systems have been widely recognised by neuroscientists, pharmacologists, and clinicians alike as key determinants (thus targets).
of a person’s mood, cognition, sleep quality, appetite, and reward systems; each of which is known to be affected by physical exercise and the cornerstone of many pharmacological treatments for major depressive disorders.

Long-term prospective data has demonstrated qualitatively distinct populations within patients with major depressive disorder, and a more recent systematic review of 67,318 participants enrolled to longitudinal cohort studies identified that people with subthreshold depression had an elevated risk of developing major depressive disorder. However, the existence for a continuum of depressive symptomology is inconclusive and subsyndromal symptoms are not yet classified as a continuum. Taken together, it remains prudent to respect the binary threshold separating clinical diagnosis with subclinical depressive symptomology, and it stands to reason that clinical and mentally healthy categories should not be ‘merged’ into the same network in order to compare head-to-head effectiveness of different exercise treatments.

With these aspects in mind, the purpose of this systematic review and network meta-analysis was to quantitatively assess the best evidence from RCTs to establish relative (head-to-head) effectiveness of resistance, aerobic, and mind-body exercise in adults aged ≥65 years below the clinical threshold for diagnosed depression. More specifically, we investigated whether (i) resistance, aerobic, and mind-body exercise training can induce substantive treatment effect on depressive symptoms in older adults, (ii) while considering relative treatment compliance to each exercise regimen, and further, (iii) to juxtapose the optimal exercise treatment for all adults aged ≥65 years irrespective of depressive symptomology.

Methods

This review was prospectively registered with PROSPERO (registration number: CRD42018115866, 23/11/2018). The network meta-analysis extension for the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA-NMA) guidelines, and the Cochrane Intervention Review that Compares Multiple Interventions provided further support in guiding this review. Guidelines specific for geriatric meta-analyses were consulted to identify baseline characteristics, equity considerations, inclusion/exclusion criteria, known confounders, and potentially important effect modifiers. Extended data for this review can be found online.

Eligibility criteria

Studies were eligible for inclusion if they (i) followed an RCT protocol, (ii) used a wait-list, usual care, or attention-control group, (iii) included one or more aerobic, resistance, or mind-body exercise intervention arms, (iv) reported depressive symptoms as an outcome at baseline and during follow-up, (v) used one or more psychometrically validated depression questionnaires, (vi) recruited participants with a minimum mean sample age of 65 years. Studies were excluded when (i) the intervention condition used a multicomponent treatment including non-exercise components with the exercise condition, or (ii) the participants were diagnosed with clinical depression, defined by DSM or ICD criteria, or a clinical threshold on a questionnaire validated against a structured diagnostic interview prior to study enrollment. Eligibility was not restricted to specific years, languages, or publication status.

Literature search

Studies were identified from computerised searches of the following databases: PubMed, Web of Science, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Health Source: Nursing/Academic Edition, PsycARTICLES, PsycINFO, and SPORTDiscus. Databases were searched for key terms pertaining to four main concepts: older adults, exercise, depressive symptoms, and RCTs. Search terms were identified using text mining procedures and an example of a search strategy is presented in detail as Extended data. Published studies, systematic reviews, and meta-analyses were also screened for additional articles. Databases were searched up to and including September 12th, 2019.

Study selection

All articles collated from the systematic search initially went through a title and abstract screening, followed by a full-text screening. Eligibility criteria were used to determine whether each article should be included or excluded. The screening process was performed concurrently for both the current review and a parallel review. Once the screening process was complete, the remaining studies were included in the systematic review and descriptively reported. Studies were only included in the quantitative analysis if they contained sufficient outcome data.

Data extraction and coding

Detailed data extraction was undertaken independently by a minimum of two researchers (KJM, PA, and/or DH) in compliance with a data extraction form (see Extended data), with any inconsistencies being arbitrated by another researcher (CM). Study characteristics are presented in Table 1. Methodological characteristics of each study were used to evaluate the quality of evidence, which included publication status, intention-to-treat principle, use of a cluster design, and validated measure(s) of depressive symptoms used.

Control groups within each individual study were further categorised as either wait-list, usual care, or attention-control. Participants undergoing wait-list conditions received the exercise intervention following trial completion. Participants randomised to usual care were those in sole receipt of conventional treatment during the trial. Participants randomised to attention-control conditions (also known as attention placebo control or active control) received activities completely unrelated to physical activity and/or exercise during their respective trials (e.g., social activities, educational programs, etc.).

Important participant and intervention characteristics were used to verify whether potential effect modifiers were similarly distributed across comparisons within the network. Participant characteristics were coded according to sample size, age (mean and standard deviation), percentage of females, and place of...
<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Treatment M (SD), females (%)</th>
<th>Control M (SD), females (%)</th>
<th>Source of participants</th>
<th>Inclusion criteria</th>
<th>Intention-to-treat</th>
<th>Cluster design</th>
<th>Treatment group</th>
<th>Control group</th>
<th>Length of intervention</th>
<th>Adherence (%)</th>
<th>Outcome measure(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adler (2007)</td>
<td>72.8 (5.4) 100</td>
<td>70.8 (8.0) 87.5</td>
<td>Community</td>
<td>Osteoarthritis</td>
<td>Yes</td>
<td>N/R</td>
<td>n = 8</td>
<td>10 weeks</td>
<td>n = 6</td>
<td>85</td>
<td>CESD-20</td>
</tr>
<tr>
<td>Angeles (2016)</td>
<td>65.0 (3.3) 80</td>
<td>69.3 (7.4) 70</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>No</td>
<td>n = 10</td>
<td>12 weeks</td>
<td>N/R</td>
<td>POMS-D</td>
<td></td>
</tr>
<tr>
<td>Ansa (2015)</td>
<td>82.8 (2.8) 65.2</td>
<td>82.6 (2.6) 65.2</td>
<td>Community</td>
<td>Sedentary</td>
<td>Yes</td>
<td>No</td>
<td>n = 23</td>
<td>16 weeks</td>
<td>N/R</td>
<td>GDS-15</td>
<td></td>
</tr>
<tr>
<td>Antunes (2005)</td>
<td>68.1 (5.5) 0</td>
<td>65.9 (3.8) 0</td>
<td>Community</td>
<td>Sedentary</td>
<td>N/R</td>
<td>No</td>
<td>n = 23</td>
<td>6 months</td>
<td>N/R</td>
<td>GDS-30</td>
<td></td>
</tr>
<tr>
<td>Bernard (2015)</td>
<td>65.5 (4.0) 100</td>
<td>65.5 (4.4) 100</td>
<td>Community</td>
<td>N/R</td>
<td>Yes</td>
<td>No</td>
<td>n = 61</td>
<td>6 months</td>
<td>53.8</td>
<td>BDI</td>
<td></td>
</tr>
<tr>
<td>Author (year)</td>
<td>Treatment M (SD), females (%)</td>
<td>Control M (SD), females (%)</td>
<td>Source of participants</td>
<td>Inclusion criteria</td>
<td>Intention-to-treat</td>
<td>Cluster design</td>
<td>Treatment group</td>
<td>Control group</td>
<td>Length of intervention</td>
<td>Adherence (%)</td>
<td>Outcome measure(s)</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>--------------</td>
<td>-----------------------</td>
<td>--------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Bethany (2005)</td>
<td>83.1 (4.8) 92.9</td>
<td>Community</td>
<td>N/R</td>
<td>No</td>
<td>n = 11 Aerobic (chair aerobics), supervised, group, N/R, 3 times/week, 30 mins/session</td>
<td>n = 10 Attention-control (social games)</td>
<td>6 weeks</td>
<td>71.7</td>
<td>BDI-II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blumenthal (1991)</td>
<td>66.5 (4.3) 48.4</td>
<td>Community</td>
<td>Sedentary</td>
<td>No</td>
<td>n = 31 Aerobic (ergometric cycling, walking, jogging, and arm ergometry), supervised, group, moderate intensity (70% HRmax), 3 times/week, 60 mins/session</td>
<td>n = 32 Wait-list</td>
<td>16 weeks</td>
<td>95.8</td>
<td>CESD-20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>67.8 (5.9) 50</td>
<td></td>
<td></td>
<td></td>
<td>n = 34 Mind-body (yoga), supervised, group, low intensity, 2 times/week, 60 mins/session</td>
<td></td>
<td></td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonura (2014)</td>
<td>77.0 (7.3) 69.8</td>
<td>Community</td>
<td>N/R</td>
<td>No</td>
<td>n = 33 Resistance (chair strength and balance), supervised, group, N/R, 1 time/week, 45 mins/session **Practice on their own for 15 minutes during non-class days</td>
<td>n = 32 Wait-list</td>
<td>6 weeks</td>
<td>83.8</td>
<td>GDS-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>n = 33 Mind-body (chair yoga), supervised, group, N/R, 1 time/week, 45 mins/session **Practice on their own for 15 minutes during non-class days</td>
<td></td>
<td></td>
<td>95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author (year)</td>
<td>Treatment M±SD, females (%)</td>
<td>Control M±SD, females (%)</td>
<td>Source of participants</td>
<td>Inclusion criteria</td>
<td>Intention-to-treat</td>
<td>Cluster design</td>
<td>Treatment group</td>
<td>Control group</td>
<td>Length of intervention</td>
<td>Adherence (%)</td>
<td>Outcome measure(s)</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------</td>
<td>---------------------------</td>
<td>-----------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>--------------</td>
<td>----------------------</td>
<td>----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Bostrom (2016)</td>
<td>84.4 (6.2) 75.3</td>
<td>85.9 (7.8) 76.3</td>
<td>Residential</td>
<td>Dementia</td>
<td>N/R</td>
<td>Yes (36)</td>
<td>n = 83</td>
<td>Resistance (functional weight-bearing and balance - High-Intensity Functional Exercise Program), supervised, group, moderate-high intensity, 2.5 times/week, 45 mins/session</td>
<td>n = 80</td>
<td>Attention-control (seated social activities)</td>
<td>4 months</td>
</tr>
<tr>
<td>Bouaziz (2019)</td>
<td>72.9 (2.5) 70</td>
<td>74.3 (3.4) 76.6</td>
<td>Community</td>
<td>Sedentary</td>
<td>N/R</td>
<td>No</td>
<td>n = 27</td>
<td>Aerobic (cycle ergometry), supervised, individual, moderate intensity, 2 times/week, 30 mins/session</td>
<td>n = 29</td>
<td>Wait-list</td>
<td>9.5 weeks</td>
</tr>
<tr>
<td>Brown (2009)</td>
<td>81.5 (6.9) 91.6</td>
<td>78.1 (6.4) 86.8</td>
<td>Community</td>
<td>N/R</td>
<td>Yes</td>
<td>Yes (8)</td>
<td>n = 26</td>
<td>Mind-body (flexibility and relaxation), supervised, group, low intensity, 2 times/week, 60 mins/session</td>
<td>n = 34</td>
<td>Usual care</td>
<td>6 months</td>
</tr>
<tr>
<td>Cancela (2016)</td>
<td>80.6 (8.3) 43.8</td>
<td>82.9 (7.4) 81</td>
<td>Residential</td>
<td>Dementia</td>
<td>Yes</td>
<td>No</td>
<td>n = 73</td>
<td>Aerobic (low resistance cycling), supervised, individual, N/R, 7 times/week, 15+ mins/session</td>
<td>n = 116</td>
<td>Attention-control (recreational activities)</td>
<td>15 months</td>
</tr>
<tr>
<td>Chen (2009)</td>
<td>65.8 (4.3) 83.9</td>
<td>72.4 (6.0) 62.1</td>
<td>Community</td>
<td>N/R</td>
<td>N/R</td>
<td>Yes (8)</td>
<td>n = 62</td>
<td>Mind-body (silver yoga), supervised, group, N/R, 3 times/week, 70 mins/session</td>
<td>n = 66</td>
<td>Wait-list</td>
<td>6 months</td>
</tr>
<tr>
<td>Chen (2015)</td>
<td>79.2 (7.0) 49.1</td>
<td>72.4 (7.1) 62.1</td>
<td>Residential</td>
<td>Wheelchair-bound</td>
<td>N/R</td>
<td>Yes (10)</td>
<td>n = 59</td>
<td>Aerobic (resistance bands, aerobic motion, and harmonic stretching - Wheelchair-bound Senior Elastic Band Program), supervised, group, N/R, 3 times/week, 40 mins/session</td>
<td>n = 55</td>
<td>Usual care</td>
<td>6 months</td>
</tr>
<tr>
<td>Author (year)</td>
<td>Treatment M-age (SD), females (%)</td>
<td>Control M-age (SD), females (%)</td>
<td>Source of participants</td>
<td>Inclusion criteria</td>
<td>Intention-to-treat</td>
<td>Cluster design</td>
<td>Treatment group</td>
<td>Control group</td>
<td>Length of intervention</td>
<td>Adherence (%)</td>
<td>Outcome measure(s)</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>-------------</td>
<td>----------------------</td>
<td>----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Chen (2017)</td>
<td>80.7 (8.0) 58.5</td>
<td>81.6 (6.7) 54.8</td>
<td>Residential</td>
<td>Dementia; wheelchair-bound</td>
<td>N/R</td>
<td>Yes (8)</td>
<td>n = 65</td>
<td>Usual care</td>
<td>15 months</td>
<td>96.9</td>
<td>CSDD</td>
</tr>
<tr>
<td>Chin A Paw (2004)</td>
<td>81.0 (5.8) 73.2</td>
<td>81.3 (4.4) 82.9</td>
<td>Residential</td>
<td>N/R</td>
<td>Yes</td>
<td>No</td>
<td>n = 41</td>
<td>Attention-control (educational programs)</td>
<td>6 months</td>
<td>78</td>
<td>GDS-30</td>
</tr>
<tr>
<td>Choi (2018)</td>
<td>77.6 (5.69) 90.9</td>
<td>78.8 (5.83) 96.7</td>
<td>Residential</td>
<td>N/R</td>
<td>N/R</td>
<td>Yes (6)</td>
<td>n = 33</td>
<td>Wait-list</td>
<td>12 weeks</td>
<td>N/R</td>
<td>GDS-15</td>
</tr>
<tr>
<td>Clegg (2014)</td>
<td>79.4 (7.9) 73.3</td>
<td>78.0 (10.5) 69.2</td>
<td>Community</td>
<td>Frailty</td>
<td>Yes</td>
<td>No</td>
<td>n = 40</td>
<td>Usual care</td>
<td>12 weeks</td>
<td>67</td>
<td>GDS-15</td>
</tr>
<tr>
<td>Collier (1997)</td>
<td>71 (N/R) 63</td>
<td>69 (N/R) 46.2</td>
<td>Community</td>
<td>N/R</td>
<td>N/R</td>
<td>No</td>
<td>n = 25</td>
<td>Usual care</td>
<td>10 weeks</td>
<td>N/R</td>
<td>GDS-30</td>
</tr>
<tr>
<td>Conradsson (2010)</td>
<td>85.3 (6.1) 73.6</td>
<td>84.2 (6.8) 72</td>
<td>Residential</td>
<td>N/R</td>
<td>Yes</td>
<td>Yes (34)</td>
<td>n = 75</td>
<td>Attention-control (seated social activities)</td>
<td>3 months</td>
<td>72</td>
<td>GDS-15</td>
</tr>
<tr>
<td>Author (year)</td>
<td>Treatment M (SD), females (%)</td>
<td>Control M (SD), females (%)</td>
<td>Source of participants</td>
<td>Inclusion criteria</td>
<td>Intention-to-treat</td>
<td>Cluster design</td>
<td>Treatment group</td>
<td>Control group</td>
<td>Length of intervention</td>
<td>Adherence (%)</td>
<td>Outcome measure(s)</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>------------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>---------------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Cramer (2016)</td>
<td>68.7 (9.1) 37</td>
<td>67.8 (10.4) 40.7</td>
<td>Community</td>
<td>Colorectal cancer</td>
<td>Yes</td>
<td>No</td>
<td>n = 27</td>
<td>Mind-body (hatha yoga), supervised, group, N/R, 1 time/week, 90 mins/session</td>
<td>n = 27 Wait-list</td>
<td>10 weeks 53</td>
<td>HADS-D</td>
</tr>
<tr>
<td>De Lima (2019)</td>
<td>66.2 (5.5) N/R</td>
<td>67.2 (5.2) N/R</td>
<td>Community</td>
<td>Parkinson's disease</td>
<td>N/R</td>
<td>No</td>
<td>n = 17</td>
<td>Resistance (free weights and machines), supervised, group, N/R, 2 times/week, 30-40 mins/session</td>
<td>n = 16 Usual care</td>
<td>20 weeks N/R</td>
<td>HRSD</td>
</tr>
<tr>
<td>Donesky-Cuenco (2009)</td>
<td>72.2 (6.5) 73.3</td>
<td>67.7 (11.5) 71.4</td>
<td>Community</td>
<td>Chronic obstructive pulmonary disease</td>
<td>Yes</td>
<td>No</td>
<td>n = 14</td>
<td>Mind-body (yoga), supervised, group, N/R, 2 times/week, 60 mins/session</td>
<td>n = 15 Usual care</td>
<td>12 weeks 83.3</td>
<td>CESD-20</td>
</tr>
<tr>
<td>Dong (2013)</td>
<td>68.5 (N/R) 100</td>
<td>N/R</td>
<td>Community</td>
<td>N/R</td>
<td>N/R</td>
<td>No</td>
<td>n = 32</td>
<td>Mind-body (YijinJing qigong), supervised, group, N/R, 3 times/week, 60 mins/session</td>
<td>N/R</td>
<td>12 weeks GDS-15</td>
<td></td>
</tr>
<tr>
<td>Dorner (2007)</td>
<td>86.7 (6.1) 76.7</td>
<td>86.9 (5.7) 76.7</td>
<td>Residential</td>
<td>Frailty; cognitive impairment</td>
<td>N/R</td>
<td>No</td>
<td>n = 15</td>
<td>Resistance (resistance bands, soft weights, and balance), supervised, group, N/R, 3 times/week, 50 mins/session</td>
<td>n = 15 N/R</td>
<td>10 weeks 91.8</td>
<td>GDS-15</td>
</tr>
<tr>
<td>Emery (1990)</td>
<td>72 (6) 83.3</td>
<td>N/R</td>
<td>Community</td>
<td>N/R</td>
<td>N/R</td>
<td>No</td>
<td>n = 14</td>
<td>Aerobic (aerobic walking and rhythmic muscle strengthening), supervised, group, moderate intensity (70% HRmax) 3 times/week, 60 mins/session</td>
<td>n = 11 Wait-list</td>
<td>12 weeks 61-94</td>
<td>CESD-20</td>
</tr>
<tr>
<td>Author (year)</td>
<td>Treatment</td>
<td>Control</td>
<td>Source of participants</td>
<td>Inclusion criteria</td>
<td>Intention-to-treat</td>
<td>Cluster design</td>
<td>Treatment group</td>
<td>Control group</td>
<td>Length of intervention</td>
<td>Adherence (%)</td>
<td>Outcome measures</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>---------</td>
<td>------------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>--------------</td>
<td>----------------------</td>
<td>--------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Eyigor (2009)</td>
<td>Aerobic (folkloric dance), supervised, group, N/R, 3 times/week, 60 mins/session</td>
<td>Usual care</td>
<td>Community</td>
<td>Cognitive impairment</td>
<td>N/R</td>
<td>N/R</td>
<td>n = 19</td>
<td>n = 18</td>
<td>8 weeks</td>
<td>N/R</td>
<td>GDS-30</td>
</tr>
<tr>
<td>Frye (2007)</td>
<td>Aerobic (walking), supervised, individual, low-moderate intensity (40-60% HR_max), 3 times/week, 20–30 mins/session</td>
<td>Usual care</td>
<td>Community</td>
<td>Diastolic heart failure</td>
<td>N/R</td>
<td>N/R</td>
<td>n = 15</td>
<td>n = 13</td>
<td>12 weeks</td>
<td>N/R</td>
<td>CES-D-20, GDS-15</td>
</tr>
<tr>
<td>Escobar (2017)</td>
<td>Mind-body (Tai Chi), supervised, group, N/R, 2 times/week, 60 mins/session</td>
<td>Usual care</td>
<td>Community</td>
<td>Sedentary</td>
<td>N/R</td>
<td>N/R</td>
<td>n = 27</td>
<td>n = 29</td>
<td>12 weeks</td>
<td>N/R</td>
<td>DASS-D</td>
</tr>
<tr>
<td>Fakhari (2017)</td>
<td>Mind-body (Tai Chi), supervised, group, N/R, 3 times/week, 20-25 mins/session</td>
<td>Usual care</td>
<td>Residential</td>
<td>Osteoarthritis</td>
<td>Yes</td>
<td>N/R</td>
<td>n = 56</td>
<td>n = 41</td>
<td>12 weeks</td>
<td>N/R</td>
<td>BD-II, GDS-30</td>
</tr>
<tr>
<td>Haque (2017)</td>
<td>Mind-body (Tai Chi), supervised, group, N/R, 3 times/week, 20-25 mins/session</td>
<td>Usual care</td>
<td>Residential</td>
<td>Sedentary</td>
<td>N/R</td>
<td>N/R</td>
<td>n = 43</td>
<td>n = 35</td>
<td>12 weeks</td>
<td>N/R</td>
<td>GDS-30</td>
</tr>
<tr>
<td>Fransen (2007)</td>
<td>Mind-body (Tai Chi), supervised, group, N/R, 3 times/week, 60 mins/session</td>
<td>Wait-list</td>
<td>Community</td>
<td>Osteoarthritis</td>
<td>Yes</td>
<td>N/R</td>
<td>n = 56</td>
<td>n = 41</td>
<td>12 weeks</td>
<td>72.2-100</td>
<td>N/R</td>
</tr>
<tr>
<td>Galley (2004)</td>
<td>Aerobic (walking), supervised, individual, low-moderate intensity (40-60% HR_max), 3 times/week, 20–30 mins/session</td>
<td>Usual care</td>
<td>Community</td>
<td>Diastolic heart failure</td>
<td>N/R</td>
<td>N/R</td>
<td>n = 15</td>
<td>n = 13</td>
<td>12 weeks</td>
<td>N/R</td>
<td>CES-D-20, GDS-15</td>
</tr>
</tbody>
</table>

**Notes:**
- **n** = sample size
- **N/R** = not reported
- **SD** = standard deviation
- **HR_max** = maximum heart rate
- **CES-D-20** = Center for Epidemiologic Studies Depression Scale-Version 20
- **GDS-15** = Geriatric Depression Scale-Version 15
- **GDS-30** = Geriatric Depression Scale-Version 30
- **BDI-II** = Beck Depression Inventory-II
- **DASS-D** = Depression Anxiety Stress Scale-Version D
- **Ta Chi** = Tai Chi
- **Walk** = Walking
- **Adverse events** = Dizziness
- **Minimum intensity** = low
- **Moderate intensity** = moderate
- **High intensity** = high
- **Memory enhancement training** = Memory enhancement training
<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Treatment M_{age} (SD), females (%)</th>
<th>Control M_{age} (SD), females (%)</th>
<th>Source of participants</th>
<th>Inclusion criteria</th>
<th>Intention-to-treat</th>
<th>Cluster design</th>
<th>Treatment group</th>
<th>Control group</th>
<th>Length of intervention</th>
<th>Adherence (%)</th>
<th>Outcome measure(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gusi (2008)</td>
<td>71 (5) 100</td>
<td>74 (6) 100</td>
<td>Community</td>
<td>Moderate depression or overweight</td>
<td>N/R</td>
<td>No</td>
<td>n = 55 Aerobic (walking), supervised, group, N/R, 3 times/week, 50 mins/session</td>
<td>n = 51 Usual care</td>
<td>6 months</td>
<td>N/R</td>
<td>GDS-15</td>
</tr>
<tr>
<td>Hars (2014)</td>
<td>75 (8) 97</td>
<td>76 (6) 95.6</td>
<td>Community</td>
<td>Risk of falling</td>
<td>Yes</td>
<td>No</td>
<td>n = 66 Aerobic (aerobic training), supervised, group, N/R, 1 time/week, 60 mins/session</td>
<td>n = 68 Wait-list</td>
<td>6 months</td>
<td>79</td>
<td>HADS-D</td>
</tr>
<tr>
<td>Hsu (2016)</td>
<td>80.7 (9.7) 63.3</td>
<td>81.7 (6.3) 63.3</td>
<td>Residential</td>
<td>Wheelchair-bound</td>
<td>Yes</td>
<td>No</td>
<td>n = 30 Mind-body (seated Tai Chi), supervised, group, N/R, 3 times/week, 40 mins/session</td>
<td>n = 30 Usual care</td>
<td>26 weeks</td>
<td>85.3</td>
<td>GDS-15</td>
</tr>
<tr>
<td>Irwin (2012)</td>
<td>70.7 (5.9) 69.6</td>
<td>71.4 (7.7) 51.4</td>
<td>Community</td>
<td>N/R</td>
<td>Yes</td>
<td>No</td>
<td>n = 58 Mind-body (Tai Chi), supervised, group, N/R, 3 times/week, 40 mins/session</td>
<td>n = 53 Attention-control (health education)</td>
<td>16 weeks</td>
<td>83</td>
<td>BDI</td>
</tr>
<tr>
<td>Irwin (2014)</td>
<td>66.3 (7.4) 64.6</td>
<td>66.4 (7.7) 72</td>
<td>Community</td>
<td>Insomnia</td>
<td>Yes</td>
<td>No</td>
<td>n = 48 Mind-body (Tai Chi), supervised, group, N/R, 1 time/week, 120 mins/session</td>
<td>n = 25 Attention-control (sleep seminar education)</td>
<td>4 months</td>
<td>81</td>
<td>IDS-C</td>
</tr>
<tr>
<td>Kekalainen</td>
<td>68.9 (2.7) 53.8</td>
<td>67.7 (2.8) 59.6</td>
<td>Community</td>
<td>N/R</td>
<td>Yes</td>
<td>No</td>
<td>n = 24 Resistance (progressive resistance training), supervised, group, N/R, 2 (month 1–3) and 1 (month 4–9) times/week, 60 mins/session</td>
<td>N/R</td>
<td>9 months</td>
<td>N/R</td>
<td>BDI-II</td>
</tr>
<tr>
<td></td>
<td>69.0 (3.3) 57.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>n = 25 Resistance (progressive resistance training), supervised, group, N/R, 2 (month 1–3) and 2 (month 4–9) times/week, 60 mins/session</td>
<td>N/R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>n = 27 Resistance (progressive resistance training), supervised, group, N/R, 2 (month 1–3) and 3 (month 4–9) times/week, 60 mins/session</td>
<td>N/R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author (year)</td>
<td>Control group</td>
<td>Treatment group</td>
<td>Cluster design</td>
<td>Intention-to-treat</td>
<td>Source of participants</td>
<td>Inclusion criteria</td>
<td>Length of intervention</td>
<td>Adherence (%)</td>
<td>Outcome measure(s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>-----------------------</td>
<td>------------------</td>
<td>-----------------------</td>
<td>---------------</td>
<td>------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim (2019)</td>
<td>Community</td>
<td>Resistence (strength exercises), N/R, N/R, moderate intensity (9-13/20 RPE), 3 times/week, 30-60 mins/session</td>
<td>N/R</td>
<td>No</td>
<td>Community</td>
<td>Reticence (strength exercises), N/R, N/R, moderate intensity (9-13/20 RPE), 3 times/week, 30-60 mins/session</td>
<td>N/R</td>
<td>24 weeks</td>
<td>GDS-15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kohut (2005)</td>
<td>Community</td>
<td>Aerobic (treadmill or ergometer, 6 times/week, 30-60 mins/session)</td>
<td>N/R</td>
<td>No</td>
<td>Sedentary</td>
<td>Aerobic (treadmill or ergometer, 6 times/week, 30-60 mins/session)</td>
<td>N/R</td>
<td>10 months</td>
<td>GDS-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krishnamurthy (2007)</td>
<td>Residential</td>
<td>Mind-body (yoga, N/R, 3 times/week, 60 mins/session)</td>
<td>N/R</td>
<td>No</td>
<td>Residential</td>
<td>Mind-body (yoga, N/R, 3 times/week, 60 mins/session)</td>
<td>N/R</td>
<td>24 weeks</td>
<td>GDS-15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuo (2018)</td>
<td>Clinical</td>
<td>Aerobic stepping exercises, N/R, 3 times/week, 30-60 mins/session</td>
<td>N/R</td>
<td>No</td>
<td>Clinical</td>
<td>Aerobic stepping exercises, N/R, 3 times/week, 30-60 mins/session</td>
<td>N/R</td>
<td>24 weeks</td>
<td>GDS-15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kupecz (2001)</td>
<td>Community</td>
<td>Aerobic (walking), N/R, supervised, group, low-moderate intensity (50-80% HRmax), 3 times/week, 45 mins/session</td>
<td>N/R</td>
<td>No</td>
<td>Community</td>
<td>Aerobic (walking), N/R, supervised, group, low-moderate intensity (50-80% HRmax), 3 times/week, 45 mins/session</td>
<td>N/R</td>
<td>24 weeks</td>
<td>CESD-20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li (2001)</td>
<td>Community</td>
<td>Mind-body (Tai Chi), supervised, group, moderate-high intensity (17-19/20 RPE), 2 times/week, 45 mins/session</td>
<td>N/R</td>
<td>No</td>
<td>Community</td>
<td>Mind-body (Tai Chi), supervised, group, moderate-high intensity (17-19/20 RPE), 2 times/week, 45 mins/session</td>
<td>N/R</td>
<td>24 weeks</td>
<td>CESD-20</td>
<td><strong>Encouraged to practice at home</strong></td>
<td></td>
</tr>
<tr>
<td>Author (year)</td>
<td>Treatment M (SD), females (%)</td>
<td>Control M (SD), females (%)</td>
<td>Source of participants</td>
<td>Inclusion criteria</td>
<td>Intention-to-treat</td>
<td>Cluster design</td>
<td>Treatment group</td>
<td>Control group</td>
<td>Length of intervention</td>
<td>Adherence (%)</td>
<td>Outcome measure(s)</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td>------------------</td>
<td>--------------------</td>
<td>---------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>----------------------</td>
<td>---------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Lin (2007)²¹</td>
<td>77.1 (7.8) 51</td>
<td>76.2 (7.3) 51</td>
<td>Community</td>
<td>History of falling</td>
<td>N/R</td>
<td>No</td>
<td>n = 39 Resistance (muscle strengthening and balance), supervised, individual, N/R, 0.5 times/week, 40–60 mins/session **Instructed to practice exercises at least three times a week</td>
<td>n = 40 Attention-control (educational programs)</td>
<td>4 months</td>
<td>N/R</td>
<td>GDS-15</td>
</tr>
<tr>
<td>Lincoln (2011)²²</td>
<td>66.0 (7.9) 69</td>
<td>66.6 (7.4) 58.6</td>
<td>Community</td>
<td>Type 2 diabetes</td>
<td>N/R</td>
<td>No</td>
<td>n = 29 Resistance (progressive resistance training), supervised, N/R, high intensity, 3 times/week, 45 mins/session</td>
<td>n = 29 Attention-control (regular phone calls)</td>
<td>16 weeks</td>
<td>N/R</td>
<td>GDS-30</td>
</tr>
<tr>
<td>Ma (2018)²³</td>
<td>70.2 (10.3) 31.7</td>
<td>69.7 (10.8) 30.4</td>
<td>Community</td>
<td>Hypertension</td>
<td>Yes</td>
<td>No</td>
<td>n = 79 Mind-body (Tai Chi), supervised, group, N/R, 4 times/week, 60 mins/session</td>
<td>n = 79 Usual care</td>
<td>6 months</td>
<td>N/R</td>
<td>CESD-10</td>
</tr>
<tr>
<td>Maki (2012)²⁴</td>
<td>71.9 (4.1) 69.3</td>
<td>72.0 (3.9) 72</td>
<td>Community</td>
<td>N/R</td>
<td>Yes</td>
<td>No</td>
<td>n = 66 Aerobic (walking and group work), supervised, group, N/R, 1 time/week, 90 mins/session</td>
<td>n = 67 Attention-control (educational programs)</td>
<td>3 months</td>
<td>87.5</td>
<td>GDS-30</td>
</tr>
<tr>
<td>Martinez (2015)²⁵</td>
<td>76.1 (8.1) 89.7</td>
<td>72.5 (8.0) 86.2</td>
<td>Clinical</td>
<td>N/R</td>
<td>Yes</td>
<td>No</td>
<td>n = 29 Mind-body (qigong), supervised, group, N/R, 2 times/week, 90 mins/session</td>
<td>n = 25 Usual care</td>
<td>4 weeks</td>
<td>88.8</td>
<td>GDS-5</td>
</tr>
<tr>
<td>Author (year)</td>
<td>Treatment M (SD), females (%)</td>
<td>Control M (SD), females (%)</td>
<td>Source of participants</td>
<td>Inclusion criteria</td>
<td>Intention-to-treat</td>
<td>Cluster design</td>
<td>Treatment group</td>
<td>Control group</td>
<td>Length of intervention</td>
<td>Adherence (%)</td>
<td>Outcome measure(s)</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>--------------</td>
<td>----------------------</td>
<td>---------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Martins (2011)</td>
<td>75.9 (7.0) 66.7</td>
<td>77.7 (8.8) 58.1</td>
<td>Residential</td>
<td>Sedentary</td>
<td>N/R</td>
<td>No</td>
<td>n = 24</td>
<td>Aerobic (rhythmic walking and stepping sequences), supervised, group, low-high intensity (40-85% HR_{max}), 3 times/week, 45 mins/session</td>
<td>n = 31</td>
<td>Usual care</td>
<td>16 weeks</td>
</tr>
<tr>
<td>McMurdo (1993)</td>
<td>82.3 (6.9) 80</td>
<td>79.3 (6.2) 80.7</td>
<td>Residential</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>n = 10</td>
<td>Resistance (strengthening exercises), supervised, group, low intensity, 2 times/week, 45 mins/session</td>
<td>n = 23</td>
<td>Attention-control (reminiscence sessions)</td>
<td>7 months</td>
</tr>
<tr>
<td>Monga (2005)</td>
<td>68 (4.2) 0</td>
<td>70.6 (5.3) 0</td>
<td>Community</td>
<td>Prostate Cancer</td>
<td>N/R</td>
<td>No</td>
<td>n = 11</td>
<td>Aerobic (treadmill walking), supervised, group, moderate intensity (65% HR_{max}), 3 times/week, 45–50 mins/session</td>
<td>n = 10</td>
<td>Usual care</td>
<td>8 weeks</td>
</tr>
<tr>
<td>Mortazavi (2012)</td>
<td>71.7 (8.2) 64.1</td>
<td>71.7 (8.2) 61.3</td>
<td>Community</td>
<td>N/R</td>
<td>N/R</td>
<td>No</td>
<td>n = 181</td>
<td>Aerobic (upper, lower, and whole body movements), supervised, group, low intensity, 2 times/week, 45 mins/session</td>
<td>n = 191</td>
<td>Attention-control (physical activity pamphlets)</td>
<td>2 months</td>
</tr>
<tr>
<td>Netz (1994)</td>
<td>70.7 (6.8) N/R</td>
<td>74.1 (8.9) N/R</td>
<td>Clinical</td>
<td>Cognitive or affective deterioration</td>
<td>N/R</td>
<td>No</td>
<td>n = 15</td>
<td>Aerobic (light callisthenic and rhythmic movements), supervised, group, low intensity, 3 times/week, 45 mins/session</td>
<td>n = 15</td>
<td>Attention-control (social intellectual activity)</td>
<td>8 weeks</td>
</tr>
<tr>
<td>Author (year)</td>
<td>Treatment M&lt;sub&gt;age&lt;/sub&gt; (SD), females (%)</td>
<td>Control M&lt;sub&gt;age&lt;/sub&gt; (SD), females (%)</td>
<td>Source of participants</td>
<td>Inclusion criteria</td>
<td>Intention-to-treat</td>
<td>Cluster design</td>
<td>Treatment group</td>
<td>Control group</td>
<td>Length of intervention</td>
<td>Adherence (%)</td>
<td>Outcome measure(s)</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Ng (2017)&lt;sup&gt;10&lt;/sup&gt;</td>
<td>70.3 (5.3) 56.2</td>
<td>70.1 (5.0) 56</td>
<td>Community</td>
<td>Frailty</td>
<td>Yes</td>
<td>No</td>
<td>n = 48</td>
<td>Resistance (functional strength and balance), supervised, group, moderate intensity, 2 times/week, 90 mins/session **Weeks 13-24 were instructed to participate at home</td>
<td>n = 50</td>
<td>Usual care</td>
<td>24 weeks</td>
</tr>
<tr>
<td>Oken (2006)&lt;sup&gt;11&lt;/sup&gt;</td>
<td>73.6 (5.1) 78.7</td>
<td>71.2 (4.4) 75</td>
<td>Community</td>
<td>Low activity (aerobic exercise less than 210 minutes per week)</td>
<td>N/R</td>
<td>No</td>
<td>n = 38</td>
<td>Aerobic (walking), supervised, group, moderate intensity (6-7/10 RPE; 70% HR&lt;sub&gt;max&lt;/sub&gt;), 1 time/week, 60 mins/session Adverse events: Hip pain (n = 1) **Exercise daily at least 5 times per week was strongly encouraged</td>
<td>n = 42</td>
<td>Wait-list</td>
<td>6 months</td>
</tr>
<tr>
<td>Park (2016)&lt;sup&gt;12&lt;/sup&gt;</td>
<td>75.3 (7.5) 75</td>
<td>Community</td>
<td>Osteoarthritis</td>
<td>N/R</td>
<td>No</td>
<td>n = 52</td>
<td>Mind-body (chair yoga), supervised, group, N/R, 2 times/week, 45 mins/session Adverse events: Groin muscle strain (n = 1) **Exercise daily at least 5 times per week was strongly encouraged</td>
<td>n = 48</td>
<td>Attention-control (health education)</td>
<td>8 weeks</td>
<td>95</td>
</tr>
<tr>
<td>Payne (2008)&lt;sup&gt;13&lt;/sup&gt;</td>
<td>64.7 (6.3) 100</td>
<td>Community</td>
<td>Breast cancer</td>
<td>N/R</td>
<td>N/R</td>
<td>n = 9</td>
<td>Aerobic (walking), unsupervised, individual, moderate intensity, 4 times/week, 20 mins/session</td>
<td>n = 9</td>
<td>Usual care</td>
<td>12 weeks</td>
<td>N/R</td>
</tr>
<tr>
<td>Pedersen (2017)&lt;sup&gt;14&lt;/sup&gt;</td>
<td>79.2 (6.6) 52.6</td>
<td>81.3 (5.1) 50</td>
<td>Residential</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>n = 19</td>
<td>Resistance (resistance training), supervised, group, N/R, 2 times/week, 60 mins/session</td>
<td>n = 12</td>
<td>N/R</td>
<td>12 weeks</td>
</tr>
<tr>
<td>Author (year)</td>
<td>Treatment</td>
<td>Control group</td>
<td>Adherence (%)</td>
<td>Length of intervention</td>
<td>Cluster design</td>
<td>Intention-to-treat criteria</td>
<td>Source of participants</td>
<td>Inclusion criteria</td>
<td>Treatment group</td>
<td>Control group</td>
<td>Outcome measures</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>--------------</td>
<td>---------------</td>
<td>------------------------</td>
<td>----------------</td>
<td>--------------------------</td>
<td>----------------------</td>
<td>------------------</td>
<td>----------------</td>
<td>--------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Penninx (2002)</td>
<td>149</td>
<td>144</td>
<td>N/R</td>
<td>18 months</td>
<td>Attention-control (health education)</td>
<td>No</td>
<td>Yes</td>
<td>Community</td>
<td>Osteoarthritis</td>
<td>Aerobic (walking), N/R, moderate intensity (50-70% HR max), 3 times/week, 60 mins/session</td>
<td>Aerobic (walking), N/R, moderate intensity (50-70% HR max), 3 times/week, 60 mins/session</td>
</tr>
<tr>
<td>Pinniger (2013)</td>
<td>8</td>
<td>9</td>
<td>N/R</td>
<td>4 weeks</td>
<td>N/R</td>
<td>No</td>
<td>Yes</td>
<td>Community</td>
<td>Age-related macular degeneration</td>
<td>Aerobic (tango dance), supervised, group, N/R, 2 times/week, 90 mins/session</td>
<td>Resistance (dumbbell and cuff weight exercises), supervised, N/R, 3 times/week, 60 mins/session</td>
</tr>
<tr>
<td>Ranganath (2017)</td>
<td>68.2 (8.8)</td>
<td>68.2 (8.8)</td>
<td>68.9 (N/R)</td>
<td>100</td>
<td>N/R</td>
<td>No</td>
<td>Yes</td>
<td>Residential</td>
<td>N/R</td>
<td>Resistance (dumbbell and cuff weight exercises), supervised, N/R, moderate intensity (50-70% HR max), 3 times/week, 60 mins/session</td>
<td>Resistance (dumbbell and cuff weight exercises), supervised, N/R, moderate intensity (50-70% HR max), 3 times/week, 60 mins/session</td>
</tr>
<tr>
<td>Rosianni (2019)</td>
<td>84.5 (4.8)</td>
<td>84.2 (6.9)</td>
<td>85.4 (4.7)</td>
<td>100</td>
<td>N/R</td>
<td>No</td>
<td>Yes</td>
<td>Residential</td>
<td>Frailty</td>
<td>Resistance (dumbbell and cuff weight exercises), supervised, N/R, high intensity (70-80% HR max), 3 times/week, 40 mins/session</td>
<td>Resistance (dumbbell and cuff weight exercises), supervised, N/R, high intensity (70-80% HR max), 3 times/week, 40 mins/session</td>
</tr>
<tr>
<td>Sahin (2018)</td>
<td>97</td>
<td>97</td>
<td>98</td>
<td>16</td>
<td>N/R</td>
<td>No</td>
<td>Yes</td>
<td>Residential</td>
<td>N/R</td>
<td>Resistance (dumbbell and cuff weight exercises), supervised, N/R, high intensity (70-80% HR max), 3 times/week, 40 mins/session</td>
<td>Resistance (dumbbell and cuff weight exercises), supervised, N/R, high intensity (70-80% HR max), 3 times/week, 40 mins/session</td>
</tr>
</tbody>
</table>

Note: N/R = Not reported; SD = Standard deviation; HR max = Heart rate maximum; BDII = Beck Depression Inventory II; GDS-15 = Geriatric Depression Scale-15; CES-D-6 = Center for Epidemiologic Studies Depression Scale-6; HRSD = Hamilton Rating Scale for Depression; BDI-II = Beck Depression Inventory II; N/R = Not reported; M = Mean; SD = Standard deviation; % = Percentage; n = Number of participants.
<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Treatment M &lt;sub&gt;age&lt;/sub&gt; (SD), females (%)</th>
<th>Control M &lt;sub&gt;age&lt;/sub&gt; (SD), females (%)</th>
<th>Source of participants</th>
<th>Inclusion criteria</th>
<th>Intention-to-treat</th>
<th>Cluster design</th>
<th>Treatment group</th>
<th>Control group</th>
<th>Length of intervention</th>
<th>Adherence (%)</th>
<th>Outcome measure(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sattin (2005)</td>
<td>80.4 (3.1) 95.4</td>
<td>80.5 (3.2) 93.6</td>
<td>Residential</td>
<td>Frailty; history of falling</td>
<td>Yes</td>
<td>Yes (20)</td>
<td>n = 158</td>
<td>Mind-body (Tai Chi), supervised, group, N/R, 2 times/week, 60-90 mins/session</td>
<td>n = 153</td>
<td>Attention-control (health education)</td>
<td>48 weeks</td>
</tr>
<tr>
<td>Sola-Serrabou (2019)</td>
<td>71.9 (5.0) 76.5</td>
<td>74.8 (6.1) 77.8</td>
<td>Residential</td>
<td>Sedentary</td>
<td>N/R</td>
<td>No</td>
<td>n = 18</td>
<td>Resistance (strength exercises), supervised, N/R, moderate intensity (5-6/10 RPE), 2 times/week, 60 mins/session</td>
<td>n = 17</td>
<td>Usual care</td>
<td>24 weeks</td>
</tr>
<tr>
<td>Song (2019)</td>
<td>76.22 (5.76) 80</td>
<td>75.33 (6.78) 70</td>
<td>Community</td>
<td>Cognitive impairment; low activity (aerobic exercise less than 150 minutes per week)</td>
<td>Yes</td>
<td>No</td>
<td>n = 60</td>
<td>Aerobic (stepping exercises), supervised, moderate intensity (12-14/20 RPE), 3 times/week, 60 mins/session</td>
<td>n = 60</td>
<td>Attention-control (health education)</td>
<td>16 weeks</td>
</tr>
<tr>
<td>Sparrow (2011)</td>
<td>70.3 (7.5) 32.7</td>
<td>71.7 (7.2) 29.4</td>
<td>Community</td>
<td>Low activity (exercise less than 20 minutes per week)</td>
<td>Yes</td>
<td>N/R</td>
<td>n = 49</td>
<td>Resistance (resistance training - Telephone-Linked Computer-based Long-term Interactive Fitness Trainer), unsupervised, N/R, N/R, 3 times/week, 60 mins/session</td>
<td>n = 51</td>
<td>Attention-control (health education)</td>
<td>12 months</td>
</tr>
<tr>
<td>Swoap (1994)</td>
<td>65.2 (4.2) 53.8</td>
<td>65.2 (4.2) 50</td>
<td>Community</td>
<td>Sedentary</td>
<td>N/R</td>
<td>No</td>
<td>n = 24</td>
<td>Aerobic (walking), supervised, group, moderate intensity (65-70% HR&lt;sub&gt;max&lt;/sub&gt;), 3 times/week, 30–65 mins/session</td>
<td>n = 11</td>
<td>Wait-list</td>
<td>26 weeks</td>
</tr>
<tr>
<td>Author (year)</td>
<td>Control group</td>
<td>Treatment group</td>
<td>Cluster design</td>
<td>Intention-to-treat criteria</td>
<td>Inclusion criteria</td>
<td>Source of participants</td>
<td>Control M (SD)</td>
<td>Treatment M (SD)</td>
<td>Treatment</td>
<td>Length of intervention</td>
<td>Adherence (%)</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>-------------------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>------------</td>
<td>----------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Tapps (2013)**</td>
<td>n = 14</td>
<td>M age: 76 (N/R)</td>
<td>No</td>
<td>N/R</td>
<td>N/R</td>
<td>NR</td>
<td>76 (N/R)</td>
<td>71.5 (10.3)</td>
<td>n = 53</td>
<td>12 weeks</td>
<td>N/R</td>
</tr>
<tr>
<td>Taylor-Piliae (2014)**</td>
<td>n = 48</td>
<td>M age: 71.5 (10.3)</td>
<td>No</td>
<td>N/R</td>
<td>N/R</td>
<td>NR</td>
<td>71.5 (10.3)</td>
<td>68.2 (10.3)</td>
<td>n = 53</td>
<td>12 weeks</td>
<td>N/R</td>
</tr>
<tr>
<td>Tsang (2013)**</td>
<td>n = 24</td>
<td>M age: 72.9 (9.5)</td>
<td>No</td>
<td>N/R</td>
<td>N/R</td>
<td>NR</td>
<td>72.9 (9.5)</td>
<td>76.3 (8.4)</td>
<td>n = 25</td>
<td>12 weeks</td>
<td>N/R</td>
</tr>
<tr>
<td></td>
<td>n = 61</td>
<td>M age: 83.3 (6.3)</td>
<td>No</td>
<td>N/R</td>
<td>N/R</td>
<td>NR</td>
<td>83.3 (6.3)</td>
<td>84.9 (6.0)</td>
<td>n = 55</td>
<td>12 weeks</td>
<td>N/R</td>
</tr>
<tr>
<td>Tsutsumi (1998)**</td>
<td>n = 61</td>
<td>M age: 68.5 (6.1)</td>
<td>No</td>
<td>N/R</td>
<td>N/R</td>
<td>NR</td>
<td>68.5 (6.1)</td>
<td>72.7</td>
<td>n = 12</td>
<td>12 weeks</td>
<td>N/R</td>
</tr>
<tr>
<td>Ohtsuka (2017)**</td>
<td>n = 24</td>
<td>M age: 72.6 (5.3)</td>
<td>No</td>
<td>N/R</td>
<td>N/R</td>
<td>NR</td>
<td>72.6 (5.3)</td>
<td>73.7 (6.3)</td>
<td>n = 33</td>
<td>12 weeks</td>
<td>N/R</td>
</tr>
<tr>
<td>Author (year)</td>
<td>Treatment M (SD), females (%)</td>
<td>Control M (SD), females (%)</td>
<td>Source of participants</td>
<td>Inclusion criteria</td>
<td>Intention-to-treat</td>
<td>Cluster design</td>
<td>Treatment group</td>
<td>Control group</td>
<td>Length of intervention</td>
<td>Adherence (%)</td>
<td>Outcome measure(s)</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------</td>
<td>-------------------------------</td>
<td>------------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>---------------</td>
<td>------------------------</td>
<td>---------------</td>
<td>--------------------</td>
</tr>
</tbody>
</table>
| Vankova (2014) | 83.4 (8.2) 91.6               | 82.9 (7.9) 92.4               | Residential            | N/R               | N/R               | No             | n = 79  
Aerobic (ballroom dance - Exercise Dance for Seniors Program), supervised, group, N/R, 1 time/week, 60 mins/session | n = 83 Wait-list | 3 months | 84.6 | GDS-15 |
| Wang (2009)   | 75.5 (9.2) 100                | 74.5 (8.1) 80                 | Community              | N/R               | N/R               | No             | n = 7  
Mind-body (yoga), supervised, group, N/R, 2 times/week, 60 mins/session | n = 10 Attention-control (social group) | 4 weeks | 78.1 | CESD-10 |
| Witham (2005) | 80 (6) 37                     | 81 (4) 54                     | Community              | Frailty; chronic heart failure | N/R               | No             | n = 36  
Aerobic (progressive aerobic seated exercise with small weights), supervised, group, N/R, 2 times/week, 20 mins/session  
** Months 4–6 were home exercise 2–3 times per week with no face-to-face contact | n = 32 Usual care | 6 months | 82.7 | HADS-D |
| Yang (2005)   | 72.6 (5.4) 68.4               | 72.7 (7.5) 90.5               | Community              | Chronic pain      | N/R               | No             | n = 19  
Mind-body (Chun Soo Energy Healing qigong), supervised, group, N/R, 2 times/week, 20 mins/session | n = 21 Wait-list | 4 weeks | N/R | POMS-D |
| Yeh (2003)    | 65 (6) 40                     | 66 (6) 40                     | Community              | Chronic obstructive pulmonary disease | Yes N/R            | No             | n = 5  
Mind-body (Tai Chi), supervised, group, N/R, 2 times/week, 60 mins/session | n = 5 Usual care | 12 weeks | 74.2 | CESD-20 |
| Zanuso (2012) | 74.3 (3.5) 50                 | 67.1 (2.0) 50                 | Community              | Sedentary         | N/R               | No             | n = 10  
Resistance (resistance training), supervised, N/R, moderate intensity, 3 times/week, 60 mins/session | n = 10 Wait-list | 12 weeks | 90.7 | POMS-D |

Note. N/R = not reported; n = Total participants included in intention-to-treat or post-treatment data analysis; HRmax = Heart rate maximum; VO2max = Maximal oxygen uptake; 1RM = One-repetition maximum; RPE = Rating of perceived exertion; BDI = Beck Depression Inventory; CESD = Center for Epidemiological Studies Depression Scale; CSDD = Cornell Scale for Depression in Dementia; DASS-D = Depression, Anxiety, and Stress Scale (depression subscale); GADS-D = Goldberg Anxiety and Depression Scale (depression subscale); GDS = Geriatric Depression Scale; GHQ-D = General Health Questionnaire (depression subscale); HADS-D = Hospital Anxiety and Depression Scale (depression subscale); HRSD = Hamilton Rating Scale for Depression; IDS-C = Inventory of Depression Symptomatology; MADRS = Montgomery-Asberg Depression Rating Scale; POMS-D = Profile of Mood States (depression subscale) PROMIS-EDD SF-8a = Patient Reported Outcome Measurement Information System - Emotional Distress and Depression Short Form-8a; TDQ = Taiwanese Depression Questionnaire.

*Outcome measure used in network meta-analysis.

**Additional unsupervised exercise component.
residence (community-dwelling, residential care, clinical setting). Relevant inclusion criteria (e.g., sedentary, dementia, etc.) were further used to assess the risk of bias from equity considerations.

Intervention characteristics were appropriately coded. Exercise types were categorised as aerobic, resistance, or mind-body exercise. Length of interventions were coded in weeks or months. Exercise intensities were coded according to ratings of perceived exertion (RPE) score, heart rate maximum (HRmax): low = 50%, moderate = 50-70%, high = 70%), maximal oxygen uptake (VO2max; low = 40%, moderate = 40-60%, high = >60%) or one-repetition maximum (1Rmax; low = <50%, moderate = 50-74%, high = >74%) or where unavailable, this was estimated according to the assessment of the original author(s). Frequency was coded as total sessions per week. Duration was coded as the average number of minutes engaging in exercise per session. Format of program was dichotomously coded as exercise in a group or individual setting. Format of supervision was dichotomously coded as supervised or unsupervised exercise. Agreement between the three researchers was 91.3%.

Risk of bias and quality assessment
Risk of bias of the included RCTs was assessed using the Cochrane Collaboration’s Tool for Assessing Risk of Bias, which were independently conducted by a minimum of two researchers (KJM, PA, and/or DH). Discrepancies were arbitrated by another co-author (CM). Appraisal for ‘other sources of bias’ were evaluated with consideration for small sample size (n < 15), low adherence (less than 80%), cluster randomisation, and inequity in the selection of the sample.

Summary of outcomes
Outcome statistics including means (M), standard deviations (SD), and sample sizes (n) for depressive symptoms were used to calculate the mean change in the primary outcome. Test statistics (i.e., t-, F-, and p-values) were used to estimate effect sizes when descriptive statistics were unavailable. Pairwise relative (head-to-head) treatment effects for depressive symptoms were estimated using Hedges’ g, which corrects for overestimation biases due to small sample sizes. Hedges’ g coefficients were interpreted according to Cohen’s conversions, whereby effects were considered small (0.2), medium (0.5), and large (0.8). Independent subgroups (e.g., males vs. females) within studies were treated as independent effect size estimates. When individual studies reported more than one post-treatment depression score for the same group of participants, only the initial post-treatment time-point was used. When studies reported depression scores on multiple outcome measures, the included depression measure was selected in compliance with clinical applicability.

Secondary outcome data were also extracted for study attrition, treatment adherence, and adverse events. Pairwise relative (head-to-head) treatment effects for study attrition were reported as odds ratios based on the pre-treatment sample size versus post-treatment dropout in the treatment and comparison conditions. Treatment adherence were reported as a percentage of total attendance in, or compliance to, the treatment condition. Adverse events were qualitatively reported according to the descriptive information provided in the transcript.

Data synthesis
One notable assumption of network meta-analysis relates to comparison group estimates, which are derived from pooling studies with homogenous between-study effect modifiers. If the distribution of an effect modifier is heterogeneous across studies within a specific comparison group, the assumption of transitivity can be violated. In the context of this network meta-analysis, we separate exercise conditions (i.e., aerobic, resistance, mind-body) and control conditions (i.e., wait-list, usual care, attention-control) to account for the between-study variation of these effect modifiers. Given the intricacies of depression severity, a forest plot was generated to depict the juxtaposition of treatment effects between (i) participants in the present network meta-analysis with depressive symptomology but not clinically diagnosed and (ii) participants in a previous network meta-analysis with clinical depression.

Risk of bias assessment identified the potential for unit-of-analysis error within studies using a cluster design for treatment allocation. Thus, sample sizes were recalculated to account for error in cluster RCTs by determining a design effect which estimates the deviation in precision of comparison group estimates, which are derived from pooling studies with homogenous between-study effect modifiers.

Data were analysed and figures were generated using STATATA/SE 15.1. Comparison-adjusted funnel plots were used to evaluate publication bias and small-study effects. Random-effects meta-regression was used to investigate modifying effects of age, gender, source of participants, length of intervention, exercise intensity, frequency, duration, and format of program, which have been identified as potential risks to transitivity in geriatric meta-analyses.

A multivariate random-effects meta-analysis was conducted using the ‘mvmeta’ command. A random-effects model assumes variance both within and between studies, explaining the heterogeneity of treatment effects. A common heterogeneity parameter was assumed across comparisons. Heterogeneity was evaluated using tau-squared (τ²), which estimates the deviation in effect sizes across the population of studies. The 95% prediction interval (PrI) was also used to estimate the true dispersion of effect within two standard deviations of the mean effect size. The significance level was p < 0.05 for all analyses.

Results
Study selection
The initial systematic search yielded a total of 3,704 citations. When duplicate studies were removed (n = 1,395), 2,309 eligible records were identified. Titles and abstracts were screened by two independent researchers (KJM and DCGB) with an agreement of 91.9%. Subsequently, 356 full-text articles were assessed for compliance with inclusion criteria and outcome data. Full-text screening was performed independently by two
researchers (KJM and PA) with an agreement of 81.8%, where discrepancies were arbitrated and resolved by consensus with a third researcher (CM). Studies using duplicate sample sets \((n = 12)\) were identified, where the most informative publication with complete data being included in the quantitative analysis. Finally, 81 studies fulfilled all inclusion criteria to be included in the systematic review. In cases of missing data, authors were emailed by the first author (KJM). If authors failed to respond following two contact efforts, and there were insufficient data to calculate effect size estimates, the study was excluded from the quantitative analysis \((n = 9)\). Additionally, three studies\(^{60-62}\) fulfilled all criteria but necessitated exclusion from the quantitative analysis due to control conditions being insufficiently defined, and the authors being unobtainable. Figure 1 outlines study selection and additional exclusion criteria.

**Characteristics of the studies**

Data from a total of 5,379 participants (2,815 treatment and 2,564 control) across 69 studies were pooled in the quantitative analysis. Each study reported pre- and post-treatment measures.

---

**Figure 1. Flowchart of screening process.**
of depression symptoms, and data from 60 of the 69 studies were obtained for post-treatment attrition. Figure 2 illustrates the network of pairwise comparisons across all studies. Aerobic, resistance, and mind-body conditions had one or more direct comparison with each of the five comparison conditions. Direct comparisons for depressive symptoms were robustly characterised between intervention and corresponding control conditions (aerobic = 28, resistance = 22, mind-body = 32), with proportions being similar for study attrition (aerobic = 24, resistance = 21, mind-body = 25).

From the 69 included studies, 43 studies enrolled young-old samples (65–75 years) and 26 studies enrolled old-old samples (>75 years). Participants were identified as either community-dwelling (n = 46), living in residential care (n = 18), clinical (n = 4), or uncategorised (n = 1). Twelve studies enrolled participants with a range of active and sedentary lifestyles. Some studies recruited older adults that were either frail (n = 5), wheelchair bound (n = 3), or identified as presenting with a risk of falling (n = 2). Participants with age-associated comorbidities were also recruited, including cardiovascular problems (n = 6), dementia (n = 3), cognitive impairment (n = 3), cancer (n = 2), and other chronic illness (n = 6).

From 79 independent exercise interventions, the majority were supervised in group exercise classes (n = 64). The remainder were either individually supervised (n = 5), unsupervised in an individual format (n = 2), supervised in an undisclosed format (n = 1), or entirely undisclosed (n = 1). The average exercise program length (weeks) was shorter for mind-body (M = 13.67, SD = 6.77) than aerobic (M = 19.76, SD = 14.44) or resistance (M = 18.00, SD = 9.07). Average weekly exercise minutes (calculated as frequency multiplied by duration) were calculated for aerobic (M = 107.78, SD = 40.88), resistance (M = 126.90, SD = 50.28), and mind-body exercise (M = 132.75, SD = 76.07). Treatment adherence (n = 44) was comparable between aerobic (M = 81.01, SD = 13.72), resistance (M = 81.17, SD = 9.00), and mind-body exercise (M = 79.34, SD = 14.73). Adverse events were observed during several of the included trials, which are presented in Table 2.

**Risk of bias and quality assessment**

Risk of bias for each study is presented comprehensively as Extended data (see Figure S1). Both blinding of participants and personnel to treatment was implausible due to the implicit nature of exercise training interventions, and thus, the remaining six criteria were used to assess the overall risk of bias within each study. Methodological quality of included studies can be considered low-to-moderate (M = 4.08/6, where low = 1, unclear, = 0.5, high = 0). Assessment of random sequence generation (selection bias), incomplete outcome data (attrition bias), and selective outcome reporting (reporting bias) may be considered adequate for most studies, whereas allocation concealment (selection bias) and blinding of outcome assessment (detection bias) were diverse. High ‘other sources of bias’ was due to low study adherence (n = 14), small sample sizes (n = 8), or both (n = 1). See Figure 3 for a summary of the risk of bias in the included studies.

![Figure 2. Network plot of comparisons for all studies included in the network meta-analysis. Line width is proportional to the number of pairwise effect size estimates and node size is proportional to the number of participants.](image-url)
Table 2. Adverse events from included studies.

<table>
<thead>
<tr>
<th>Exercise group</th>
<th>Study</th>
<th>Adverse event(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aerobic</td>
<td>Kupecz (2001)</td>
<td>Minor medical attention (n = 2)</td>
</tr>
<tr>
<td></td>
<td>Oken (2006)</td>
<td>Hip pain (n = 1)</td>
</tr>
<tr>
<td>Resistance</td>
<td>Ansai (2015)</td>
<td>Mild muscle pain (n = 9)</td>
</tr>
<tr>
<td></td>
<td>Chin A Paw (2004)</td>
<td>Program was too intensive (n = 8)</td>
</tr>
<tr>
<td></td>
<td>Clegg (2014)</td>
<td>Fell at least once (n = 7)</td>
</tr>
<tr>
<td>Mind-body</td>
<td>Cramer (2016)</td>
<td>Muscle soreness (n = 3), abdominal pain (n = 1), neck pain (n = 1), minor vertigo (n = 1), hip pain (n = 1)</td>
</tr>
<tr>
<td></td>
<td>Eyre (2017)</td>
<td>Dizziness (n = 1)</td>
</tr>
<tr>
<td></td>
<td>Oken (2006)</td>
<td>Groin muscle strain (n = 1)</td>
</tr>
</tbody>
</table>

Figure 3. Risk of bias chart of studies included in the quantitative analysis.

Assessment of inconsistency
Inconsistency network models were used to test the global consistency of direct and indirect effects of pairwise and multiam comparisons. Assumption of consistency was satisfied for each treatment (p > 0.05). Tests for inconsistency between direct and indirect estimates were not significant (p > 0.05), thus indirect and direct estimates were not different to direct evidence. Inconsistency tables can be found as Extended data (see Tables S1 and S2).

Loop-specific heterogeneity was explored using inconsistency plots (see Extended data, Figure S7 and S8). Within the depressive symptoms network, inconsistency factors (IF) did not indicate high inconsistency (IF = 0.00 to 0.65) or loop-specific heterogeneity ($\tau^2 = 0.05$ to 0.28). The AER-MB-UC loop departed from the minimum lower-bound confidence interval (CI), yet fell short of presenting risk for heterogeneity ($IF = 0.54$, $\tau^2 = 0.05$, CI = 0.04, 1.03). Within the attrition network, the ratio of two odds ratios (RoR) indicated a high degree of inconsistency for the AER-RES-UC (RoR = 1.87) and MB-RES-UC (RoR = 1.77) loops. Each were interpreted as presenting elevated risk for heterogeneity, which were subsequently downgraded during GRADE assessment. All remaining loops satisfied assumption of consistency.

Publication bias and sensitivity analyses
Comparison-adjusted funnel plots were used to detect publication bias and small-study effects. Funnel plots were roughly symmetrical for both depressive symptoms and attrition, indicative of low risk of publication bias and no presence of small-study effects. See Figure 4 for the depressive symptoms network and Extended data for the attrition network (see Figure S9).

In order to test transitivity across networks, potential effect modifiers were tested with meta-regression sensitivity analyses for the entire pool of studies and separately for each exercise...
comparison (aerobic vs. resistance vs. mind-body). No significant modifying effects were observed for the pool of studies or any separate exercise comparison for age, gender, source of participants, length of intervention, format of program, exercise intensity, frequency, duration, adherence, year of study, risk of bias, publication status, intention-to-treat analysis, nor cluster design, indicating that the assumption of transitivity was upheld. Full analyses are presented as Extended data (see Tables S6–S8).

### Results of the network meta-analysis

Data pooled from 69 eligible studies provided a total of 88 individual comparisons for depressive symptoms and 76 individual comparisons for study attrition. Table 3 presents the network meta-analysis of depressive symptoms and attrition. Network estimates were calculated to establish relative effectiveness between pairs of comparisons.

Each exercise type effectively reduced depressive symptoms compared with control conditions (see Figure 5). Ranking of treatments for depressive symptoms are presented on SUCRA plots of ranked mean values, which can be found as Extended data (see Figure S11 and Table S3). Ranked order of quantitative values determined mind-body exercise to be the most effective type of exercise to mitigate depressive symptoms, followed closely by resistance and aerobic exercise, respectively. The magnitude of study effect did not reach statistical threshold to favour any individual exercise treatment.

Resistance exercise demonstrated the highest study compliance compared with each of the other comparison groups, but the dispersion of effect estimates presented a level of heterogeneity that confounded any substantive differences. Comprehensive reporting of study attrition can be found as Extended data (see Figure S10) in addition to SUCRA plots and ranked order (see Figure S12 and Table S4).

### Effectiveness vs. attrition

A two-dimensional clustered ranking plot was employed to illustrate the average reduction in depressive symptoms for each comparison, relative to average attrition rate. Figure 6 presents the ranking of the exercise conditions with respective control conditions in conjunction with SUCRA values for depressive symptoms (effectiveness) and attrition. The three exercise conditions amalgamated a single cluster and were more effective than control conditions.
Table 3. League table for head-to-head comparisons.

<table>
<thead>
<tr>
<th></th>
<th>Wait-list</th>
<th>Usual care</th>
<th>Resistance</th>
<th>Mind-body</th>
<th>Aerobic</th>
<th>Attention-control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.11</td>
<td>-0.34</td>
<td>-0.06</td>
<td>0.12</td>
<td>1.03</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>(-0.40, 0.17)</td>
<td>(-0.57, -0.10)</td>
<td>(-0.33, 0.22)</td>
<td>(-0.12, 0.35)</td>
<td>(0.48, 1.77)</td>
<td>(0.41, 2.57)</td>
</tr>
<tr>
<td></td>
<td>1.10</td>
<td>0.92</td>
<td>1.35</td>
<td>0.12</td>
<td>0.03</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.48, 2.48)</td>
<td>(0.47, 1.77)</td>
<td>(0.61, 2.99)</td>
<td>(-0.12, 0.35)</td>
<td>(0.50, 2.01)</td>
<td>(0.41, 2.57)</td>
</tr>
<tr>
<td></td>
<td>0.74</td>
<td>0.68</td>
<td>-0.06</td>
<td>0.06</td>
<td>0.33</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>(0.30, 1.83)</td>
<td>(0.37, 1.26)</td>
<td>(-0.33, 0.22)</td>
<td>(0.06, 0.35)</td>
<td>(0.05, 0.61)</td>
<td>(0.02, 0.52)</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.94</td>
<td>1.39</td>
<td>1.03</td>
<td>0.33</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.50, 2.01)</td>
<td>(0.48, 1.84)</td>
<td>(0.62, 3.10)</td>
<td>(0.51, 2.07)</td>
<td>(0.05, 0.61)</td>
<td>(0.02, 0.52)</td>
</tr>
<tr>
<td></td>
<td>0.74</td>
<td>0.68</td>
<td>-0.06</td>
<td>0.06</td>
<td>0.33</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>(0.30, 1.83)</td>
<td>(0.37, 1.26)</td>
<td>(-0.33, 0.22)</td>
<td>(0.06, 0.35)</td>
<td>(0.05, 0.61)</td>
<td>(0.02, 0.52)</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.94</td>
<td>1.39</td>
<td>1.03</td>
<td>0.33</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.50, 2.01)</td>
<td>(0.48, 1.84)</td>
<td>(0.62, 3.10)</td>
<td>(0.51, 2.07)</td>
<td>(0.05, 0.61)</td>
<td>(0.02, 0.52)</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.94</td>
<td>1.39</td>
<td>1.03</td>
<td>0.33</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.50, 2.01)</td>
<td>(0.48, 1.84)</td>
<td>(0.62, 3.10)</td>
<td>(0.51, 2.07)</td>
<td>(0.05, 0.61)</td>
<td>(0.02, 0.52)</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.94</td>
<td>1.39</td>
<td>1.03</td>
<td>0.33</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.50, 2.01)</td>
<td>(0.48, 1.84)</td>
<td>(0.62, 3.10)</td>
<td>(0.51, 2.07)</td>
<td>(0.05, 0.61)</td>
<td>(0.02, 0.52)</td>
</tr>
</tbody>
</table>

Note. Depressive symptoms (upper) are reported as Hedges’ g and 95% confidence intervals. Negative scores indicate a greater decrease in depressive symptoms for the column group. Attrition (lower) is reported as odds ratios and 95% confidence intervals. An odds ratio lower than 1 indicates a greater attrition for the column-defining comparison.

Figure 5. Predictive interval plot for the depressive symptoms network. Black diamonds represent the difference in the effect size estimate (Hedges’ g). Narrow horizontal lines represent the confidence intervals (CI) and the wider horizontal lines represent the prediction intervals (PrI). The blue vertical line represents the null hypothesis (Hedges’ g = 0). Negative scores indicate a greater decrease in depressive symptoms for the comparison (left) group. AC = attention-control; AER = aerobic; MB = mind-body; RES = resistance; UC = usual care; WL = wait-list.
GRADE assessment
The certainty of evidence was assessed with the GRADE approach\(^\text{129}\). All control comparisons in the depressive symptoms network were rated as high or moderate certainty, which were downgraded due to small sample size or potential risks of bias (i.e., attrition bias, detection bias, or bias resulting from low adherence). Comparisons between exercise interventions were moderate-to-low certainty, resulting from imprecision in confidence intervals and small sample sizes. Detailed summary of the depressive symptoms network is presented in Table 4. Estimates of the attrition network were downgraded due to inconsistency and imprecision, which reflect moderate to very low confidence this outcome (see Extended data, Table S5).

Generalisation for all adults aged $\geq 65$ years
Figure 7 presents collective representation for all adults aged $\geq 65$ years. This comparison employs scaled distribution (Hedges’ $g$) for the present data and that of clinically depressed older adults\(^\text{23}\), the effectiveness of aerobic, resistance, and mind-body exercise is comparably consistent for all adults aged $\geq 65$ years, irrespective of depression severity. These findings should provide reassurance for personnel and stakeholders in healthy ageing to encourage exercise prescription from a point of pragmatism and in collaboration with patient preference.

Theoretical implications for the current findings
Behavioural and physiological research\(^\text{24,25,27,28}\) coalesce to provide ample reasoning to separate participant cohorts with existing clinical depression from those without diagnosis. This study is no different. In agreement with findings from previous research\(^\text{13,14,35}\), aerobic and resistance exercise demonstrated similar treatment effectiveness in older samples aged $\geq 65$ years (Hedges’ $g = -0.06$, $PrI = -0.91$, 0.79). Exercise characteristics (i.e., intensity, frequency, duration, etc.) are often similar between aerobic and resistance exercise, representing two sides of the same coin. Meta-analytical data\(^\text{11}\) on older adults with existing clinical depression observed that exercise programs incorporating a combination of aerobic and resistance training were most beneficial. Although the synergistic effects of combined exercise types were beyond the scope of the current review, it is conceivable that aerobic and resistance exercise may complement one another in an exercise intervention.

Discussion
The present review offers new information for general exercise prescription to support mental health outcomes for adults aged $\geq 65$ years. Specifically, (i) aerobic, resistance, and mind-body each demonstrate equivalent benefit to mitigate symptoms of depression in adults aged $\geq 65$ years, (ii) compliance to exercise treatment is notably encouraging for each exercise types, and (iii) when combined with the pool of data from clinically depressed older adults\(^\text{23}\), the effectiveness of aerobic, resistance, and mind-body exercise is comparably consistent for all adults aged $\geq 65$ years, irrespective of depression severity. These findings should provide reassurance for personnel and stakeholders in healthy ageing to encourage exercise prescription from a point of pragmatism and in collaboration with patient preference.

Pooled direct and indirect estimates marginally favoured treatment with mind-body exercise over either aerobic...
Table 4. Summary of GRADE assessment for the certainty in depressive symptoms estimates.

<table>
<thead>
<tr>
<th>Comparison Effect</th>
<th>Number of participants</th>
<th>Number of direct comparisons</th>
<th>Nature of evidence</th>
<th>Certainty</th>
<th>Reason for downgrading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aerobic vs. wait-list</td>
<td>371 vs. 357</td>
<td>11</td>
<td>Mixed</td>
<td>High</td>
<td>None</td>
</tr>
<tr>
<td>Aerobic vs. usual care</td>
<td>309 vs. 310</td>
<td>9</td>
<td>Mixed</td>
<td>Moderate</td>
<td>Risk of bias&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Aerobic vs. attention-control</td>
<td>431 vs. 472</td>
<td>8</td>
<td>Mixed</td>
<td>High</td>
<td>None</td>
</tr>
<tr>
<td>Resistance vs. wait-list</td>
<td>43 vs. 42</td>
<td>2</td>
<td>Mixed</td>
<td>Moderate</td>
<td>Imprecision&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Resistance vs. usual care</td>
<td>358 vs. 272</td>
<td>15</td>
<td>Mixed</td>
<td>High</td>
<td>None</td>
</tr>
<tr>
<td>Resistance vs. attention-control</td>
<td>267 vs. 274</td>
<td>5</td>
<td>Mixed</td>
<td>Moderate</td>
<td>Risk of bias&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mind-body vs. wait-list</td>
<td>380 vs. 363</td>
<td>12</td>
<td>Mixed</td>
<td>High</td>
<td>None</td>
</tr>
<tr>
<td>Mind-body vs. usual care</td>
<td>358 vs. 356</td>
<td>10</td>
<td>Mixed</td>
<td>Moderate</td>
<td>Risk of bias&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mind-body vs. attention-control</td>
<td>298 vs. 265</td>
<td>10</td>
<td>Mixed</td>
<td>High</td>
<td>None</td>
</tr>
<tr>
<td>Aerobic vs. resistance</td>
<td>24 vs. 23</td>
<td>1</td>
<td>Mixed</td>
<td>Low</td>
<td>Imprecision&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>Aerobic vs. mind-body</td>
<td>90 vs. 83</td>
<td>4</td>
<td>Mixed</td>
<td>Moderate</td>
<td>Imprecision&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
<tr>
<td>Resistance vs. mind-body</td>
<td>33 vs. 33</td>
<td>1</td>
<td>Mixed</td>
<td>Low</td>
<td>Imprecision&lt;sup&gt;g&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>Potential attrition bias due to high number of studies with incomplete outcome data.
<sup>b</sup>Small sample size.
<sup>c</sup>Potential risk of bias due to high number of studies with low adherence.
<sup>d</sup>Potential detection bias due to high number of studies without blinding of outcome assessment.
<sup>e</sup>Confidence intervals include values favouring either treatment.

(Hedges’ g = -0.12, PrI = -0.95, 0.72) or resistance exercise (Hedges’ g = -0.06, PrI = -0.90, 0.79). However, it must be noted that the magnitude of effect falls short of being statistically different between groups and should be considered equivalent. Certainty of evidence is moderate, due to the dispersion in effect size estimates resulting in imprecision. Direct comparisons from multi-arm RCTs have offered mind-body exercise to be more effective than aerobic<sup>91</sup> or resistance<sup>44</sup> exercise. Moreover, subgroup analyses<sup>46</sup> have indicated that clinically depressed older adults respond more favourably to mind-body exercise, but this hypothesis has not be substantiated<sup>41</sup>.

Since mind-body exercise engages low intensity muscular activity (i.e., yoga, Tai Chi, qigong), the novel evidence in the current systematic analysis challenge the idea that intensity is the primary mechanism for the antidepressive effect of exercise. Rather, mind-body exercise combines the mental and physical aspects of exercise, which may result in similar antidepressive effects to higher intensity exercise<sup>134,135</sup>. Critical to these mental aspects is interoceptive sensations such as an internally directed focus on breathing and proprioception, which have previously been linked to the resilience of depressive states<sup>136,137</sup>. Thus, it is plausible that mind-body exercise allows older adults to regulate negative mood states, which is not normally possible during aerobic and resistance activities.

Other important determinants of successful programming include study attrition, adherence rate, and adverse events. Here, we hypothesised that each exercise condition would demonstrate lower compliance to treatment than wait-list, usual care, and attention-control comparisons. Contrary to expectations, pooled direct and indirect estimates indicated that study attrition was comparable for all comparisons apart from resistance exercise, which offered a higher degree of compliance. However, on deeper scrutiny of absolute sample size, any differences observed in attrition become abrogated due to relatively smaller participant numbers within the resistance exercise studies (n = 705) compared with aerobic (n = 1,143) or mind-body (n = 1,005). Thus, in consideration of the wide dispersion in effect size estimates and a moderate risk of bias in individual studies, there are limitations in the certainty to confidently conclude substantive difference in study attrition between any comparisons.

Each of the three types of exercise had similar adherence rates and time spent per week (calculated as frequency multiplied by duration). Interestingly, mind-body exercise had relatively shorter intervention length than either aerobic or resistance exercise (M<sub>aerobic</sub> = 6.09 and 4.33 weeks, respectively) despite having a greater reduction in depressive symptoms. This could be explained by (i) mind-body exercise having the same antidepressive effects in a shorter time than aerobic and resistance exercise, or (ii) a potential
plateau effect whereby the antidepressive effects reach a maximum threshold during the first 10–15 weeks of an exercise program and are then maintained during the remaining weeks. Either way, it seems plausible that mind-body exercise provides a slightly more effective intervention against depressive symptoms in older populations without clinical depression, but that these treatment effects are not substantive enough to constitute statistical difference.

**Practical considerations**

With consideration to projected population estimates over the next decade and the consequential demand on healthcare services, the findings from this network meta-analysis offer a message of support for exercise prescription to promote mentally healthy ageing. When considering the collective findings of the present review in conjunction with the recent network meta-analysis in clinically depressed adults aged ≥ 65 years, stakeholders in healthy ageing and exercise prescription have encouraging pooled RCT evidence for the antidepressant effects of either aerobic, resistance, or mind-body exercise for older adults across the mental health continuum.

Treatment safety is a matter of ongoing importance in gerontological health, and exercise treatment programs are no different. Systematic scrutiny of the included RCTs found that study participants reported no major adverse events and only a few minor somatic complaints (n = 28). Taken together, this provides encouraging support for personnel wanting to safely prescribe exercise-based intervention programs in older populations. Of course, there is always a possibility of underreporting adverse events in clinical trials, and the present review was no exception. Therefore, the importance of reporting event outcomes, adverse

![Forest plot depicting juxtaposed collective RCT evidence of older adults (n = 5,975) in the depressive symptoms network.](image-url)

**Figure 7. Forest plot depicting juxtaposed collective RCT evidence of older adults (n = 5,975) in the depressive symptoms network.** Circles represent the difference in the effect size estimate (Hedges’ g) for those with depressive symptomology but not diagnosed with clinical depression (n = 5,379). Triangles represent the difference in the effect size estimate (Hedges’ g) for those with clinical depression (n = 596; effect sizes adapted from Miller et al., 2020). Horizontal lines represent the confidence intervals (CI). The dotted vertical line represents the null hypothesis (Hedges’ g = 0). Negative scores indicate a greater decrease in depressive symptoms for the left group, and vice versa.
or otherwise, cannot be understated. In fact, there is a known phenomenon in geriatric exercise research whereby adverse events are often underreported because authors do not consider minor adverse events to be noteworthy and/or essential to the primary purpose of the trial, giving rise to an ongoing issue that will not be corrected until all studies routinely report event outcomes.

Nevertheless, participants engaging in aerobic exercise reported the least adverse events (n = 3), including minor medical attention and hip pain. Amongst studies included in this meta-analysis, aerobic exercise predominantly involved walking and stationary cycling, which may reflect a safe and natural form of exercise for older adults. Resistance exercise was typically associated with participants experiencing mild muscular pain and falls (n = 16), which may be explained by the progressive overloading of resistance-based training. Notably, incidents of falling were reported in an unsupervised exercise program. Finally, mind-body exercise was typically associated with different types of muscular pain and body strain (n = 9). It is speculated that the higher rates of injury in mind-body exercise are predominantly because it incorporates flexibility, balance, and stability movements, which may be unique to older bodies. In general, exercise seems to be a relatively safe intervention for older adults living in both the local community and residential aged care, although intensity and supervision, particularly for resistance training, should be monitored to ensure falls and injury do not occur.

The present review has some notable advantages above a traditional pairwise meta-analysis. RCTs with considerable non-exercising components, such as those using a multicomponent exercise intervention, were excluded because they may have overestimated the magnitude of the true effect in past reviews. Specifically, it is likely that multicomponent exercise interventions such as laughter therapy, depression awareness training, or self-efficacy training may have introduced a risk of bias by inflating the observed effectiveness of the exercise program on depressive symptoms through a secondary, complementary treatment effect. Pairwise meta-analysis also assumes that all control groups are the same, which is not always the case. To manage heterogeneity from this assumption, control groups were separated into individual network comparisons. Taken together, the current findings provide a more accurate estimate of the true effects of exercise on depressive symptoms in adults aged ≥ 65 years.

Limitations and future directions
The present network meta-analysis is not without limitations. As study participants and personnel cannot be successfully blinded, there is an inherent risk of performance bias. It is also believed that many exercise-based interventions have a small number of participants, shorter follow-up, and do not adequately conceal randomisation, which are all likely to reduce the quality of RCTs and increase the risk of bias. However, we mitigated the impact of this by comparing relative effects with multiple control groups in order to increase reliability and specificity. This, combined with the relatively low risk of bias in individual RCTs, were extremely important in minimising overall risk of bias and achieving accurate effect comparisons in the present review.

Since most RCTs did not explicitly describe the exclusion of participants with ongoing diagnosis of clinical depression, there was potential contamination with data from participants with existing clinical diagnosis and medical treatment that went unreported. This was primarily managed by separating (i) participants with depressive symptomology but not clinically diagnosed in the present review from (ii) participants with clinical depression in a previous network meta-analysis. Within this review, we further mitigated this effect modifier by only including RCTs, where this risk would be balanced by control participants. We recommended that ageing researchers encourage the reporting of all ongoing pharmacological regimens in trials recruiting older participants.

Although modifying effects were explored using meta-regression, potential confounding effects from exercise modifiers (e.g., fitness improvements, length of program, session frequency and duration, exercise intensity, supervision, group format) were outside the scope of our network meta-analysis. There has been a modicum of such exploration in subgroup and meta-regression analyses of previous reviews, providing researchers with an encouraging opportunity in their planning of future similar work. Future meta-analyses with extensive subgroup analyses should explain the heterogeneity of effect sizes between similar exercise intervention studies in older persons.

Conclusions
Pooled RCT evidence highlights that each individual exercise mode (aerobic, resistance, and mind-body) demonstrate equivalence to mitigate symptoms of depression in older adults, irrespective of depression severity. As each exercise treatment demonstrated encouraging levels of treatment compliance, we endorse personnel and stakeholders in healthy ageing to encourage individual/patient preference when prescribing exercise to older adults ≥ 65 years presenting with depressive symptomology.

Data availability
Underlying data
Figshare: Aerobic, resistance, and mind-body exercise are equivalent to mitigate symptoms of depression in older adults: A systematic review and network meta-analysis of randomised controlled trials (extended data). https://doi.org/10.6084/m9.figshare.12998549.v2

This project contains the following underlying data:
- Data File D1 (PDF file containing raw outcome data)
- STATA/SE 15.1 files (DTA files containing raw data)

Extended data
Figshare: Aerobic, resistance, and mind-body exercise are equivalent to mitigate symptoms of depression in older adults: A

This project contains the following extended data:
- Supplementary File S1 (PDF file containing additional information, tables, and figures not in the main manuscript)

Reporting guidelines

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Acknowledgments
The authors gratefully acknowledge the skilled contributions of Dr. Beyon Miloyan for his technical guidance, thoughtful discussions, and expert critique of the present work.

References


125. Smarr KL, Keefer AL: Measures of depression and depressive symptoms: Beck Depression Inventory-II (BDI-II), Center for Epidemiologic Studies Depression Scale (CES-D), Geriatric Depression Scale (GDS), Hospital Anxiety and Depression Scale (HADS), and Patient Health Questionnaire-9 (PHQ-9). Arthritis Care Res (Hoboken). 2011; 63 Suppl 11: S454-66. PubMed Abstract | Publisher Full Text


130. StataCorp: Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC, 2017.


Open Peer Review

Current Peer Review Status:  ?  ?  ✓

Version 1

Reviewer Report 09 June 2021

https://doi.org/10.5256/f1000research.29961.r83975

© 2021 Imboden C. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Christian Imboden
1 Private Clinic Wyss AG, Sierre, Switzerland
2 Psychiatric Services Solothurn, University of Basel, Basel, Switzerland

General
The review and network meta-analysis are very well written and conclusive. I especially appreciate the aim of the authors to analyze studies of different modalities of physical activity (PA) on depressive symptoms in older adults without clinical depression and comparing them to their already published network-MA on the effects of PA in clinically depressed older adults. The methods are described thoroughly. The authors analyze RCTs evaluating the effects of aerobic, resistance or mind-body exercise on depressive symptoms separately. In addition, they also group control groups by waitlist, attention control and care as usual which yields a very detailed analysis. They found significant positive effects on depressive symptoms for all three modalities of PA versus control interventions with no specific modality yielding significantly stronger effects than the other. They also analyzed attrition to the study protocol, an important variable that adds additional weight to the analysis. Considering attrition, the authors found comparable values in all comparisons. The results are displayed in appropriate ways by graphical and numerical means. And thus, provides a good overview of the various results. The statistical methods and interpretation seem absolutely appropriate to me, but since I am not a specialist for network meta-analysis I can not provide an in-depth evaluation.

I have some minor comments to the authors:

P 24, Results of the network-MA:
Please spell out “SUCRA” at its first mentioning

Figure 7:
In the notes, the confidence interval is not specified. Is it 95% CI? Please specify. I also don't understand why the horizontal lines (95% CI) are different from Figure 5: i.e., the comparison AER vs. WL has a 95% CI of -0.63 to -0.14. So why does the horizontal line depicting the CI cross the vertical zero-line in Figure 7? Please explain.
Discussion:
p. 27: you described that mind-body exercise has additional mental effects such as interoceptive sensations. Often mind-body exercise also incorporates aspects of mindfulness. Therefore, I believe it would be appropriate to insert a short paragraph discussing the effects of mindfulness (such as meditation and MBSR) on depressive symptoms. A suggested reference might be: Li et al. (2019) DOI: 10.1111/inm.12568

On the same page, you discuss compliance to treatment. You state, that you hypothesized that compliance would be lower in exercise treatments. However, the meta-analysis on dropout rates of Stubbs et al. (2016), DOI: 10.1016/j.jad.2015.10.019 found similar dropout rates for exercise and control conditions but analyzed samples of adult patients (not specifically elderly). Nonetheless, you might discuss this in a sentence.

Conclusions:
You state that exercise mitigates symptoms of depression regardless of depression severity. If interpreted strictly, according to the current paper this only applies to subclinical severity of symptoms. If you incorporate your earlier paper (network MA of exercise in clinically depressed elderly) into the conclusion, I believe this should be mentioned accordingly.

References

Are the rationale for, and objectives of, the Systematic Review clearly stated?
Yes

Are sufficient details of the methods and analysis provided to allow replication by others?
Yes

Is the statistical analysis and its interpretation appropriate?
I cannot comment. A qualified statistician is required.

Are the conclusions drawn adequately supported by the results presented in the review?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Depression, physical activity for depression

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
Kyle Miller, Federation University, Australia

General
The review and network meta-analysis are very well written and conclusive. I especially appreciate the aim of the authors to analyze studies of different modalities of physical activity (PA) on depressive symptoms in older adults without clinical depression and comparing them to their already published network-MA on the effects of PA in clinically depressed older adults. The methods are described thoroughly. The authors analyze RCTs evaluating the effects of aerobic, resistance or mind-body exercise on depressive symptoms separately. In addition, they also group control groups by waitlist, attention control and care as usual which yields a very detailed analysis. They found significant positive effects on depressive symptoms for all three modalities of PA versus control interventions with no specific modality yielding significantly stronger effects than the other. They also analyzed attrition to the study protocol, an important variable that adds additional weight to the analysis. Considering attrition, the authors found comparable values in all comparisons. The results are displayed in appropriate ways by graphical and numerical means. And thus, provides a good overview of the various results. The statistical methods and interpretation seem absolutely appropriate to me, but since I am not a specialist for network meta-analysis I can not provide an in-depth evaluation.

We wish to firstly thank the Reviewer for their thoughtful commentary. We have taken time to consider all feedback and make the necessary revisions to the article.

I have some minor comments to the authors:
P 24, Results of the network-MA:
Please spell out “SUCRA” at its first mentioning
We have included “surface under the cumulative ranking curve” before the first mention of the acronym.

Figure 7:
In the notes, the confidence interval is not specified. Is it 95% CI? Please specify.
We thankful to the Reviewer for picking up this clerical error, on our behalf. This has now been rectified to state the “95% CI” in notes for Figures 5 and 7.

I also don't understand why the horizontal lines (95% CI) are different from Figure 5: i.e., the comparison AER vs. WL has a 95% CI of -0.63 to -0.14. So why does the horizontal line depicting the CI cross the vertical zero-line in Figure 7? Please explain.
We are grateful to the Reviewer for highlighting a lack of consistency in the article, specifically for intervals between Figure 5 and 7. Version 1 of the article as read by Reviewers contained both confidence and prediction intervals for Figures 5 and 7, respectively. In light of this, Figure 7 is now converted to 95% confidence intervals in order to maintain observational consistency with Figure 5. We have also proofed consistency of data (and presentation thereof) within article figures.

Discussion:
p. 27: you described that mind-body exercise has additional mental effects such as interoceptive sensations. Often mind-body exercise also incorporates aspects of
mindfulness. Therefore, I believe it would be appropriate to insert a short paragraph discussing the effects of mindfulness (such as meditation and MBSR) on depressive symptoms. A suggested reference might be: Li et al. (2019) DOI: 10.1111/inm.125681

This is a very reasonable query relating to the potential for ‘mindfulness’ being an adjunct to potential improvement from mind-body exercise. The authors have taken the opportunity to record these findings within the article, with enhanced detail within the online supplementary material. This supplementary material extends upon the current article discussion to include potential discrete mechanisms that may underpin the psycho-adaptive response to exercise in older adults. This can be found within the online repository which accompanies this article:

- ‘Data Availability’ > ‘Extended data’ > ‘Supplementary File S1’
- https://doi.org/10.6084/m9.figshare.12998549.v3

On the same page, you discuss compliance to treatment. You state, that you hypothesized that compliance would be lower in exercise treatments. However, the meta-analysis on dropout rates of Stubbs et al. (2016), DOI: 10.1016/j.jad.2015.10.0192 found similar dropout rates for exercise and control conditions but analyzed samples of adult patients (not specifically elderly). Nonetheless, you might discuss this in a sentence.

This is an interesting point and one which we credited with consideration during the planning of this review. As the Reviewer will be aware, every epidemiological study conducted to date identify older adults to be the least active and least compliant with exercise compared with younger demographics. For reasons of being conservative and containing the potential for observer bias, this was handled in the negative unless data demonstrated otherwise. However, on reflection we deem the use of ‘hypothesised’ as potentially inappropriate compared with alternatives such as ‘anticipated’. This has been amended in the Discussion section to read: “Here, we anticipated that each exercise condition would demonstrate lower compliance to treatment than wait-list, usual care, and attention-control comparisons.”

Conclusions:
You state that exercise mitigates symptoms of depression regardless of depression severity. If interpreted strictly, according to the current paper this only applies to subclinical severity of symptoms. If you incorporate your earlier paper (network MA of exercise in clinically depressed elderly) into the conclusion, I believe this should be mentioned accordingly.

We extend our gratitude to this Reviewer, who in coalition with Reviewer 2, identified this very same point. On reflection, this was clearly an overreach by the authors. We fully accept the requirement for amendment, which resulted in amendments to the Abstract, Discussion, and Conclusion sections.

Competing Interests: None
Andrea Camaz Deslandes
Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

Objective:
The study aimed to investigate the head-to-head effectiveness of aerobic, resistance, and mind-body exercise in ameliorating depressive symptoms in older adults. The study is relevant and original, considering the increased prevalence of depression in older adults and the need for treatment through complementary and alternative practices, such as exercise. This topic has indeed received little attention in literature until now. The method used is adequate to the purpose of the study, and the results and the discussion were developed appropriately. However, some minor adjustments are necessary for a better understanding of the manuscript. Please find below some of the specific items regarding the issues.

Introduction:
I suggest a presentation of the concepts of physical activity and exercise, as well as a discussion about divergences in the classification of Yoga and other mind-body practices as physical exercise.

Methods:
The inclusion of older adults with Dementia (n=3), Cognitive Decline (n=3), Parkinson's Disease (n=1), and Cognitive frailty (n=1) can be a possible limitation of the present study, considering the differences related to comorbidity. I suggest the author exclude these studies. Neuropsychiatric symptoms in MCI and Dementia are related to several different mechanisms, as well as the psychological and physiological effects of physical exercise in these populations. The characteristics of Depression in these cases can be very specific, with causes, neurophysiological changes and treatments that differ from depressive symptoms in healthy older adults. The inclusion of these studies can contribute of shrinking the effect size of exercise. For example, the effectiveness of pharmacological approaches to managing depressive symptoms in Dementia and MCI is less effective and the antidepressant effect of exercise in these populations is still controversial. If the authors decide to include these population (MCI, Parkinson, Dementia), several other studies that investigated the effect of exercise on neuropsychiatric symptoms (such as depression) in patients with MCI and Dementia should be included in the present meta-analysis, For more details, see Liang 2018 and Leng et al., 2018. Moreover, this issue should be included in the limitation section.

Discussion:
The author should discuss the possible different mechanisms related to the antidepressant effect of mind-body practices, aerobic and resistance training in older adults, considering both psychological and physiological hypotheses. Please consider more recent neurobiological mechanisms related to the effect of exercise on depressive symptoms (e.g., myokines, neurotrophins, inflammatory cytokines, etc.).
Conclusion:
The authors concluded: “Pooled RCT evidence highlights that each individual exercise mode (aerobic, resistance, and mind-body) can effectively mitigate symptoms of depression in older adults, irrespective of depression severity”. I suggest excluding the phrase “irrespective of depression severity”, since the authors did not run a specific analysis to evaluate this issue.

References

Are the rationale for, and objectives of, the Systematic Review clearly stated?
Yes

Are sufficient details of the methods and analysis provided to allow replication by others?
Yes

Is the statistical analysis and its interpretation appropriate?
Yes

Are the conclusions drawn adequately supported by the results presented in the review?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Physical exercise and mental health

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.
opportunity to improve our article. The author team have taken time to deliberate on each point and have reached consensus within our responses below. Thank you.

**Introduction:**

*I suggest a presentation of the concepts of physical activity and exercise, as well as a discussion about divergences in the classification of Yoga and other mind-body practices as physical exercise.*

We acknowledge this as an important point, and one which would require detailed expansion and integration of broad knowledge of the different forms of exercise classification and their nuanced differences. In order to preserve the flow of rationale within the Introduction section, while similarly being complete with exercise characteristics of aerobic, resistance, and mind-body regimens, we have provided supplementary reading to satisfy the broader readership of exercise specialists, including background material for readers out-with this specialist topic.

We have taken this opportunity to include further statement within the Introduction section to highlight the differences between types of exercise: “We direct the reader to the Extended data of this article for detailed description of the exercise classifications included in this review.”

‘Extended data’ for publication in F1000 is repository for non-essential information or data. To diligently tackle important divergences in classifications in exercise type required 750 words, which the authors deemed as an important point with an appropriate home within the online repository, which can be found here:

- ‘Data Availability’ > ‘Extended data’ > ‘Supplementary File S1’
- https://doi.org/10.6084/m9.figshare.12998549.v3

**Methods:**

*The inclusion of older adults with Dementia (n=3), Cognitive Decline (n=3), Parkinson’s Disease (n=1), and Cognitive frailty (n=1) can be a possible limitation of the present study, considering the differences related to comorbidity. I suggest the author exclude these studies. Neuropsychiatric symptoms in MCI and Dementia are related to several different mechanisms, as well as the psychological and physiological effects of physical exercise in these populations. The characteristics of Depression in these cases can be very specific, with causes, neurophysiological changes and treatments that differ from depressive symptoms in healthy older adults. The inclusion of these studies can contribute of shrinking the effect size of exercise. For example, the effectiveness of pharmacological approaches to managing depressive symptoms in Dementia and MCI is less effective and the antidepressant effect of exercise in these populations is still controversial.*

We agree with the Reviewers comments relating to the antidepressant effects of exercise, and in particular that evidence in these populations remains controversial. Before undertaking this review, specific inclusion criteria for this work were considered greatly, wherein, the authors remained agnostic to any particular effect size resulting from exercise participation on depressive symptoms in older adults. The purpose of the present review was to provide point-specific information that is generalisable to all older adults aged >65 years without diagnosis of clinical depression.

The binary approach taken within this review was (a) without diagnosis of clinical depression, thus excluding (b) those diagnosed with clinical depression. This review
included RCTs limited to the former, which became the defining characteristic. Specifically, we cannot bias the inclusion criteria based on ubiquitous conditions, and as such, any given adult aged >65 years is by definition a primary candidate for each of the suggested limiting covariates.

We are genuinely grateful to this Reviewer for generating some important points relating to exercise gerontology which remain unresolved. Until such time as epidemiological controversy such as nuanced and unresolved characteristics amongst ageing cohorts is evidenced, it is our belief that we must accept perceived imperfections within the literature, which is a notable limitation of meta-analytical investigation. This Reviewer does imply an important fact, and we do not offer this article as being definitive as such potential limitations are unresolved. However, we do believe that it is best practice to present in the article’s current form, but importantly, accept and acknowledge the value of the Reviewer’s observation by noting this as a generalised limitation within the limitations section. By doing this, there is opportunity for researchers to improve on this research, should new evidence add confirmatory knowledge. Supplementary sensitivity analysis specific to this can be found in the online repository, located here:

- ‘Data Availability’ > ‘Extended data’ > ‘Additional Sensitivity Analysis’
- https://doi.org/10.6084/m9.figshare.12998549.v3

If the authors decide to include these population (MCI, Parkinson, Dementia), several other studies that investigated the effect of exercise on neuropsychiatric symptoms (such as depression) in patients with MCI and Dementia should be included in the present meta-analysis, For more details, see Liang 2018 and Leng et al., 2018. Moreover, this issue should be included in the limitation section.

We are further grateful to this Reviewer for providing a list of potential references and have applied article inclusion criteria in review of each. From this list, one study (i.e., Abd El-Kader & Al-Jiffri, 2016) met the inclusion criteria for the systematic review but was not eligible for inclusion in network meta-analytics. Thus, the Systematic Review component of this article has been updated, yet the network meta-analysis remains unchanged. In response to the Reviewer’s recommendation, we have carefully examined the review conducted by Leng et al. (2018). We have applied the present article inclusion criteria in review of each study and further included detailed reasoning for inclusion/exclusion for each which may be observed at:

- ‘Data Availability’ > ‘Extended data’ > ‘Suggested Articles Provided During Peer Review’
- https://doi.org/10.6084/m9.figshare.12998549.v3

Factors related to the prescription of exercise (frequency, intensity, duration, length, supervision, cognitive engagement) should be more explored in analyses and results. This is an important point to consider (and partially addressed in the previous comment). One primary difference between network meta-analysis and its more widely used counterpart, the ‘meta-analysis’, relates to the limited ability for network meta-analysis to address significant covariates by sub-group analytics and meta-regression of known moderators of the primary outcome. However, such important factors have a place to address this within network meta-analysis. Although not overtly observable, we have done so within the present article. More specifically, “in order to test transitivity across networks, potential effect modifiers were tested with meta-regression sensitivity analyses for the
Discussion:
The author should discuss the possible different mechanisms related to the antidepressant effect of mind-body practices, aerobic and resistance training in older adults, considering both psychological and physiological hypotheses. Please consider more recent neurobiological mechanisms related to the effect of exercise on depressive symptoms (e.g., myokines, neurotrophins, inflammatory cytokines, etc.). Possible different mechanisms, and unique characteristics relating to the antidepressant potential of different forms of exercise, remain to be fully defined. The main purpose of meta-analytical review of randomised controlled trials (RCTs) is to leverage the work of other researchers, and by doing so improve statistical power and point estimation for a specific question. In this respect, the present review cannot offer psychological or physiological hypotheses to a degree greater than any one specific RCT. However, we do accept this Reviewer’s guidance and further accept that it is important to contextualise contemporary mechanisms relating to the antidepressant effect of exercise. In light of this, we have included a separate extended piece of 784 words within the online repository to provide an extension of this discussion, found here:
- ‘Data Availability’ > ‘Extended data’ > ‘Supplementary File S1’
- https://doi.org/10.6084/m9.figshare.12998549.v3

Additionally, the authors have added the following text “we direct the reader to the Extended data for further discussion relating to potential mechanisms for the antidepressant effects of exercise in older adults”, to add value to this overall body of work and contemporary mechanistic information for the interested reader.

Conclusion:
The authors concluded: “Pooled RCT evidence highlights that each individual exercise mode (aerobic, resistance, and mind-body) can effectively mitigate symptoms of depression in older adults, irrespective of depression severity”. I suggest excluding the phrase “irrespective of depression severity”, since the authors did not run a specific analysis to evaluate this issue.

Thank you for drawing our attention to this critical oversight from the authors. The authors agree and this has now been removed for reasons of specific accuracy. Specifically, this has been moderated and subsequently amended within the Abstract, Discussion, and Conclusion sections.

**Competing Interests:** None

Reviewer Report 08 March 2021

https://doi.org/10.5256/f1000research.29961.r80099

© 2021 Abdelbasset W. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Walid Kamal Abdelbasset**

1 Department of Health and Rehabilitation Sciences, College of Applied Medical Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia
2 Department of Physical Therapy, Kasr Al-Aini Hospital, Cairo University, Cairo, Egypt

This study aims to examine the head-to-head effectiveness of aerobic, resistance, and mind-body exercise to mitigate depressive symptoms in adults aged ≥ 65 years. The topic is interesting, however, some corrections should be addressed.

**General comments:**
- The English presentation needs to be improved throughout the manuscript.
- This study is a systematic review. Therefore, transfer kindly the PRISMA checklist, which can help clarify the missed information.

**Specific comments:**

**Introduction:**
- Define depression in detail. Determine the relationship between exercise and depression status in older adults. I suggest the authors expose the following points in the introduction: What is known about depression and exercise training in older adults? What is not known? Why the study was done? The hypothesis is also missed in this section.

**Discussion:**
- The main findings should be presented in the introductory paragraph with respect to the study objectives and hypothesis.
- Inconsistency models need more details in the discussion section.
- Please, re-frame the discussion along the following lines:
Main findings of the present study
Comparison with other studies
Implication and explanation of findings
Strengths and limitations
Conclusion, recommendation, and future direction.

Are the rationale for, and objectives of, the Systematic Review clearly stated?
Partly

Are sufficient details of the methods and analysis provided to allow replication by others?
Yes

Is the statistical analysis and its interpretation appropriate?
Yes

Are the conclusions drawn adequately supported by the results presented in the review?
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Exercise Intervention; Geriatric Rehabilitation; Rehabilitation Sciences

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

---

**Author Response 30 Jun 2021**

**Kyle Miller**, Federation University, Australia

This study aims to examine the head-to-head effectiveness of aerobic, resistance, and mind-body exercise to mitigate depressive symptoms in adults aged $\geq 65$ years. The topic is interesting, however, some corrections should be addressed.

We extend our gratitude to the Reviewer for their critique and offering an opportunity to improve our article. The author team have taken the time to reflect on each comment and reached consensus, which are provided in detail below.

**General comments:**

- **The English presentation needs to be improved throughout the manuscript.**

The majority of the authorship team are native English language speakers and while each of us acknowledge that we are in continual development, we are experienced in communicating scientific writing in our native language. However, we do acknowledge this criticism in the spirit of improvement that it is meant. Thus, we have taken time to consider literary communication and addressed to the best of our collective ability.

- **This study is a systematic review. Therefore, transfer kindly the PRISMA checklist, which can help clarify the missed information.**
Should we be accurate in the interpretation of this comment, the Reviewer considers this article to be (i) a systematic review. This is correct. However, it is a systematic review combined with (ii) network meta-analysis, which requires compliance with a specific set of reporting guidelines that are an extension of the noted PRISMA checklist. This article is presented in acknowledgement of a complimentary extension to PRISMA guidelines where we have adhered to the ‘PRISMA-NMA guidelines’, where ‘NMA’ accounts for a quantitative addition within this ‘network meta-analysis’. To support our response of confirmatory compliance, we direct the Reviewer to the Methodology section within this article which specifies a formal statement of compliance with PRISMA-NMA guidelines. To further offer confirmatory compliance with PRISMA-NMA guidelines, we direct this Reviewer to supplementary materials within the online repository, found here:

- ‘Data Availability’ à ‘Reporting Guidelines’ à ‘PRISMA-NMA Checklist’
- https://doi.org/10.6084/m9.figshare.12998549.v3

**Specific comments:**

**Introduction:**

**Define depression in detail.**

We thank the Reviewer for identifying potentially important rationale for conducting this systematic review and network meta-analysis. In order to provide a comprehensive response to the query, it is necessary to offer two components of ‘depression’:

- Definition of depression – This is an interesting request, and the authors believe it to be adequately rationalised within the Introduction section. In compliance PRISMA-NMA criteria, we have specified the generalised term ‘depression’ as having a characteristic binary format for this work where explicitly (i) ‘clinically defined depression’ is divergent with (ii) ‘depressive symptoms’ which fall short of the threshold for the former. The present article focuses on the latter, where we offer contextual dialogue of literature within the domain of ‘clinical depression’ and articulate within the Discussion section of this article.

- Operational definition for this article – The discrete definition for depressive symptoms used in this review are further detailed within the methodology, where RCTs were deemed eligible where valid determination of depressive symptoms was achieved using accepted psychometrically valid tools.

**Determine the relationship between exercise and depression status in older adults. I suggest the authors expose the following points in the introduction:**

- **What is known about depression and exercise training in older adults? What is not known?**

We direct the Reviewer to Paragraphs 2 of the Introduction section: “International public health consortia are in concert with the antidepressant effects of exercise as a low-risk adjunct for optimal mental health (American Psychiatric Association, 2010; Stubbs et al., 2018; World Health Organization, 2010). While we do not yet have the answers for low uptake of exercise in older adults (Knowles et al., 2015), it may in some way be due to nuanced regimen design, which in turn, may similarly impact compliance in exercise prescription by primary and stakeholders in aged care.”

Similarly, Paragraph 3 of the same Introduction section: “During the past four decades, widely different metabolic, social, and environmental demands between exercise modalities...
(i.e., running vs. weightlifting vs. Tai Chi) have been well-characterised. Given that there is variation between exercise regimens, and these variations are not merely semantics, one may be surprised to discover that only a few randomised controlled trials (RCTs) have deliberately compared the antidepressant effects of different exercise regimens in older adults (Martins et al., 2011; Penninx et al., 2002).”  

The authors acknowledge that may not be ideal for the Reviewer and are optimistic that the body of referenced material in coalition with the short steps of logical reasoning, offer mitigatory satisfaction of this Reviewer’s exposed points.

○ **Why the study was done?**

We agree that this is an important attribute which must be present within empirically sound research efforts. Accordingly, we direct the Reviewer to view the final paragraph within the Introduction section, which reads: “The purpose of this systematic review and network meta-analysis was to quantitively assess the best evidence from RCTs to establish relative (head-to-head) effectiveness of resistance, aerobic, and mind-body exercise in adults aged ≥ 65 years, and below the clinical threshold for diagnosed depression. More specifically, we investigated whether (i) resistance, aerobic, and mind-body exercise training can induce substantive treatment effect on depressive symptoms in older adults, (ii) while considering relative treatment compliance to each exercise regimen, and further, (iii) to juxtapose the optimal exercise treatment for all adults aged ≥ 65 years irrespective of depressive symptomology.”

○ **The hypothesis is also missed in this section.**

We are thankful to the Reviewer for addressing this important consideration. PRIMA-NMA guidelines stipulate that those wishing to conduct meta-analytical review should “provide an explicit statement of questions being addressed, with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).”  

There is opportunity to investigate this further, where we direct the Reviewer the emergent work from Bruce Charlton titled ‘The used and abuses of meta-analysis’, wherein it is prescribed that “meta-analysis is not a hypothesis-testing activity” and where Charlton further offers that meta-analysis “cannot legitimately be used to establish the reality of a putative hazard or therapy” (Charlton, 1996). In this respect, we prefer to present the research questions in its current active format.  

More generally, we offer that the majority of the author team are experimental scientists where research is beset on the presentation of formal hypotheses. We agree that it could be a route to detail a specific question of focus. However, as (for example) the present work relies upon and leverages the efforts from many RCT authors to satisfy our independent question, we must also acknowledge that their individual nuanced differences become amplified to induce collective heterogeneity towards reaching a resolution to our research question. For reasons such as this, we selected to side with formal guidelines (in this case, PRISMA-NMA) to achieve the outcomes detailed within the present article.

**Discussion:**

○ **The main findings should be presented in the introductory paragraph with respect to the study objectives and hypothesis.**

The authors offer some satisfaction to resolving the ‘hypothesis’ component of the
Reviewer’s suggestion in the prior response. In terms of addressing the substantive part of this Reviewer’s query, we offer the main findings from this review to be presented at the beginning of the Discussion section, and specifically, in the inaugural paragraph, which reads: “The present review offers new information for general exercise prescription to support mental health outcomes for adults aged ≥ 65 years. Specifically, (i) aerobic, resistance, and mind-body each demonstrate equivalent benefit to mitigate symptoms of depression in adults aged ≥ 65 years, (ii) compliance to exercise treatment is notably encouraging for each exercise types, and (iii) when combined with the pool of data from clinically depressed older adults the effectiveness of aerobic, resistance, and mind-body exercise is comparably consistent for all adults aged ≥ 65 years, irrespective of depression severity. These findings should provide reassurance for personnel and stakeholders in healthy ageing to encourage exercise prescription from a point of pragmatism and in collaboration with patient preference.”

- **Inconsistency models need more details in the discussion section.**

Within the present article, inconsistency models are detailed within the Results section (see ‘Assessment of Inconsistency’). Our generalised ambition in delivering coherence within the Discussion section was to integrate our findings from inconsistency modelling and to communicate in a more active voice in appreciative context with other available literature. As it turned out, our tests of inconsistency became largely satisfied with generalised assumption of consistency for (i) global and (ii) loop-specific heterogeneity for our main (depressive symptom) network which limited an expanded discussion of inconsistency. To support this Reviewer’s query and in our compliance with Cochrane Review guidelines (Higgins et al., 2021), network inconsistency was further interrogated using the GRADE assessment tool (Salanti et al., 2014). In doing this, any important inconsistencies are dialogued in context to the primary objective, which in turn, are dialogued within the Discussion section and benchmarked against our primary outcomes. For instance, Paragraph 5 of the Discussion section reads: “Certainty of evidence is moderate, due to the dispersion in effect size estimates resulting in imprecision.”

**Please, re-frame the discussion along the following lines:**
- **Main findings of the present study**
- **Comparison with other studies**
- **Implication and explanation of findings**
- **Strengths and limitations**
- **Conclusion, recommendation, and future direction.**

The authors have partially satisfied this request (RE: ‘Specific Comments - Discussion’), with requisite response above. In compliance with previously identified guidelines, we have attempted to provide a coherent Discussion section in this article which is presented in the following generalised flow:
- Paragraph 1: Summary of main findings.
- Paragraph 2-6: Theoretical implications and comparisons with past studies.
- Paragraph 7-10: Practical considerations and recommendations.
Paragraph 14: Conclusion and final recommendation.

We thank this Reviewer for constructive critique and the authors are satisfied that we have an improved article with support from Reviewer commentary.

References


Competing Interests: None
The benefits of publishing with F1000Research:

• Your article is published within days, with no editorial bias
• You can publish traditional articles, null/negative results, case reports, data notes and more
• The peer review process is transparent and collaborative
• Your article is indexed in PubMed after passing peer review
• Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com