Effectiveness of physical activity interventions for improving depression, anxiety and distress: an overview of systematic reviews

Ben Singh, Timothy Olds, Rachel Curtis, Dorothea Dumuid, Rosa Virgara, Amanda Watson, Kimberley Szeto, Edward O’Connor, Ty Ferguson, Emily Eglitis, Aaron Miatke, Catherine EM Simpson, Carol Maher

ABSTRACT

Objective To synthesise the evidence on the effects of physical activity on symptoms of depression, anxiety and psychological distress in adult populations.

Design Umbrella review.

Data sources Twelve electronic databases were searched for eligible studies published from inception to 1 January 2022.

Eligibility criteria for selecting studies Systematic reviews with meta-analyses of randomised controlled trials designed to increase physical activity in an adult population and that assessed depression, anxiety or psychological distress were eligible. Study selection was undertaken in duplicate by two independent reviewers.

Results Ninety-seven reviews (1039 trials and 128 119 participants) were included. Populations included healthy adults, people with mental health disorders and people with various chronic diseases. Most reviews (n=77) had a critically low A Measuremment Tool to Assess systematic Reviews score. Physical activity had medium effects on depression (median effect size=−0.43, IQR=−0.66 to −0.27), anxiety (median effect size=−0.42, IQR=−0.66 to −0.26) and psychological distress (effect size=−0.60, 95% CI −0.78 to −0.42), compared with usual care across all populations. The largest benefits were seen in people with depression, HIV and kidney disease, in pregnant and postpartum women, and in healthy individuals. Higher intensity physical activity was associated with greater improvements in symptoms. Effectiveness of physical activity interventions diminished with longer duration interventions.

Conclusion and relevance Physical activity is highly beneficial for improving symptoms of depression, anxiety and distress across a wide range of adult populations, including the general population, people with diagnosed mental health disorders and people with chronic disease. Physical activity should be a mainstay approach in the management of depression, anxiety and psychological distress.

PROSPERO registration number CRD42021292710.

INTRODUCTION

Mental health disorders are among the leading causes of the global health-related burden, with substantial individual and societal costs. In 2019, one in eight people (970 million) worldwide were affected by a mental health disorder and almost one in two (44%) will experience a mental health disorder in their lifetime. The annual global costs of mental health disorders have been estimated at $2.5 trillion (USD), which is projected to increase to $6 trillion (USD) by 2030. Depression is the leading cause of mental health-related disease burden, while anxiety is the most prevalent mental health disorder. Additionally, the COVID-19 pandemic has been associated with increased rates of psychological distress, with prevalence ranging between 35% and 38% worldwide.

The role of lifestyle management approaches, such as exercise, sleep hygiene and a healthy diet, varies between clinical practice guidelines in different countries. In US clinical guidelines, psychotherapy or pharmacotherapy is recommended as the initial treatment approaches, with lifestyle approaches considered as ‘complementary alternative treatments’ where psychotherapy and pharmacotherapy are ‘ineffective or unacceptable’. In other countries such as Australia, lifestyle management is recommended as the first-line treatment approach, though in practice, pharmacotherapy is often provided first.

There have been hundreds of research studies examining the effects of physical activity (PA) on depression, anxiety and psychological distress, many of which suggest that PA may have similar effects to psychotherapy and pharmacotherapy (and with numerous advantages over psychotherapy and pharmacotherapy, in terms of cost, side-effects and ancillary health benefits). Despite the evidence for the benefits of PA, it has not been widely adopted therapeutically. Patient resistance, the difficulty of prescribing and monitoring PA in clinical settings, as well as the huge volume of largely incommensurable studies, have probably impeded a wider take-up in practice.

Meta-reviews are systematic reviews of systematic reviews, offering a way of synthesising a vast evidence base. While there have been several meta-reviews of PA for depression, anxiety and psychological distress, they have focused on specific population subgroups, particular conditions (eg, depression only) or on particular forms of PA. We set out to undertake the most comprehensive synthesis to date of evidence regarding the effects of all modes of PA on symptoms of depression, anxiety and psychological distress in adult populations.
Review

Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

Selection criteria and search strategy
The population, intervention, comparison, outcomes and study type (PICOS) framework was used to develop the inclusion criteria as follows: population: any adult population (aged ≥18 years); intervention: interventions designed to increase PA. The following definition of PA was used: ‘any bodily movement produced by the contraction of skeletal muscles that results in a substantial increase in caloric requirements over resting energy expenditure’.26 Reviews were eligible irrespective of PA modality, supervision, delivery (eg, in-person or online) or dose (frequency, intensity and duration). Reviews were ineligible if they included any randomised control trials (RCTs) of non-PA interventions, if PA was combined with another intervention (eg, diet) or if they evaluated single bouts of acute exercise. Comparator: reviews were eligible if ≥75% of the included RCTs involved either usual care, waitlist, nothing an equal attention intervention or a lower/lesser PA intervention (eg, a supervised exercise intervention vs printed PA materials). During study selection, it became apparent that the comparator inclusion/exclusion criteria needed elaboration. After careful consideration and discussion, we decided to exclude reviews where >25% of component RCTs compared PA to pharmaceutical interventions or compared two types of equal dose exercise (eg, resistance vs aerobic exercise) without a non-PA comparison, since the inclusion of such reviews would limit our ability to evaluate the effectiveness of PA per se. Outcomes: any self-report or clinician-rated assessment of depression, anxiety or psychological distress symptoms. Study type: systematic reviews with meta-analyses of RCTs only, which included meta-analyses of the outcomes of interest.

Twelve databases were searched (CINAHL, Cochrane, Embase, MEDLINE, Emcare, ProQuest Health and Medical Complete, ProQuest Nursing and Allied Health Source, PsycINFO, Scopus, Sport Discus, EBSCOhost and Web of Science) using subject heading, keyword and Medical Subject Headings (MeSH) term searches for ‘systematic review’, ‘meta-analysis’, ‘physical activity’, ‘exercise’, ‘anxiety’, ‘depression’ and ‘psychological distress’ (see online supplemental eTable 1 for the full search text).

Data management and extraction
Search results were imported into EndNote V.x9 (Clarivate, Philadelphia) where duplicates were removed, then exported into Covidence (Veritas Health Innovation, Melbourne, Australia). Title/abstract and full-text screening, data extraction and risk of bias scoring were completed in duplicate by two independent reviewers (BS and AM, CEMS, DD, EE, EO, KS, RC, RV or TF), with disagreements resolved by team discussion.

Data were extracted in duplicate by two independent reviewers (BS and AM, CEMS, DD, EE, EO, KS, RC, RV or TF) using a standardised extraction form,27 28 and discrepancies were resolved by team discussion. The risk of bias of the included reviews was assessed by two independent reviewers (BS and AM, CEMS, DD, EE, EO, KS, RC, RV or TF) in duplicate using the A MeaSurement Tool to Assess systematic Reviews (AMSTAR-2) tool.29 The AMSTAR-2 tool involves 16 items, with each item scored as yes, partial yes or no. Seven items are considered ‘critical’ and nine ‘non-critical’.29 The critical domains are protocol registration, adequacy of search strategy, justification for excluding individual studies, risk of bias assessment, appropriateness of meta-analysis methods, use of risk of bias during interpretation and assessment of publication bias. Reviews were rated as ‘high confidence’ (0 critical weakness and <3 non-critical weaknesses), ‘moderate’ (one critical weakness and ≤3 non-critical weaknesses), ‘low’ (>1 critical weakness and <3 non-critical weaknesses) or ‘critically low’ (>1 critical weakness and ≥3 non-critical weaknesses).29

Umbrella review synthesis methods
The overlap in component RCTs that were included across all eligible reviews was assessed using the Corrected Covered Area (CCA) method.30 A CCA of 100% indicates that every review included in our umbrella review comprised the same component RCTs, while a CCA of 0% indicates that every review in our umbrella review included entirely unique RCTs. The following cut-offs were used to quantify the CCA: 0%–5%='slight overlap'; 6%–10%='moderate'; 11%–15%='high' and >15%='very high' overlap.30 Publication bias was assessed by creating a funnel plot and observing the presence of asymmetries or missing sections.

Meta-analysis results from each review were presented using forest plots. Separate forest plots were created for meta-analyses reporting standardised (eg, standardised mean difference, SMD) and unstandardised effect sizes (eg, mean difference). For meta-analyses that reported standardised effect sizes, we undertook subgroup analyses for clinical status and intervention characteristics. Meta-analysis results were summarised using medians and IQRs.

The Oxford Centre for Evidence-Based Medicine levels of evidence and grades for recommendations32 were used to classify the overall level of evidence as grade A: consistent level 1 studies (ie, systematic reviews of RCTs or individual RCTs); B: consistent level 2 (ie, systematic reviews of cohort studies or individual cohort studies) or level 3 studies (ie, systematic reviews of case-control studies or individual case–control studies) or extrapolations from level 1 studies; C: level 4 studies (ie, case series) or extrapolations from level 2 or 3 studies or D: level 5 (ie, expert opinion without explicit critical appraisal) evidence or troublingly inconsistent or inconclusive studies of any level.

RESULTS
Of the 1280 records identified, 97 were eligible. They included 1039 unique (component) RCTs and the CCA was 0.66%, indicating slight overlap (see online supplemental eFigure 1 for PRISMA flowchart, including reasons for exclusions). Evaluation of funnel plots indicated no evidence of publication bias.

An overview of all reviews’ characteristics is shown in online supplemental eTable 2. There was a total of >128 119 participants (n=19 did not report the number of participants). Mean participant age ranged from 29 to 86 (median=55) years, and most reviews (n=83, 86%) involved female and male participants. An overview of all populations and PA modalities is shown in table 1. Fifteen reviews specifically involved individuals with depression33–41 and three involved individuals with anxiety.42–44 Most reviews involved various PA modes (n=70) and most (n=77) had a critically low AMSTAR-2 score (low: n=10; high: n=10, online supplemental eTable 3).

Meta-analysis results: depression
Results from 72 meta-analyses based on SMD (n=875 component RCTs, >62 040 participants) showed a medium effect in
favour of PA for reducing depression and depressive symptoms (median SMD = −0.43, IQR = −0.66 to −0.27, figure 1).

MD effect size for each instrument was: profile of mood states: −7.68 (1 review), Beck Depression Inventory: −5.53 (IQR = −6.24 to −4.81), The Edinburgh Postnatal Depression Scale: −2.97 (IQR = −3.49 to −2.44), self-rating scale: −3.99 (one review), Centre for Epidemiological Studies Depression: −0.36 (IQR = −1.25 to 0.02), Montgomery-Asberg Depression Rating Scale: −1.80 and Hospital Anxiety and Depression Scale: −1.26 (IQR = −1.41 to −1.18, online supplemental eFigure 3 and online supplemental eTable 4).

Grade of recommendation: (A) Consistent level 1 studies.

Anxiety
Results from 28 meta-analyses using SMD (171 component RCTs, >10 952 participants) showed a medium effect of PA for reducing anxiety (median SMD = −0.42, IQR = −0.66 to −0.26, figure 2).

MD effect sizes for each instrument were: The State-Trait Anxiety Inventory: −4.57 (1 review), Hospital Anxiety and Depression Scale: −1.26 (IQR = −1.26 to –0.79, online supplemental eTable 4 and online supplemental eFigure 5).

Grade of recommendation: (A) Consistent level 1 studies.

Psychological distress
One systematic review reported SMD results for psychological distress (six component RCTs, 508 participants), while another systematic review reported MD results (one component RCT, 39 participants). Results showed a medium effect in favour of PA, compared with usual care (SMD = −0.60, 95% CI −0.78 to −0.42). For MD, findings showed no significant effect (MD = −0.30, 95% CI −5.55, 4.95, one review, one component RCT, 39 participants).

Grade of recommendation: (B) Consistent level 2 or 3 studies or extrapolations from level 1 studies.

### Table 1 Overview of all populations, conditions and physical activity modes of the included reviews

<table>
<thead>
<tr>
<th>Population or condition</th>
<th>Number of reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>n=27</td>
</tr>
<tr>
<td>Adults with depression</td>
<td>n=17</td>
</tr>
<tr>
<td>Dementia</td>
<td>n=5</td>
</tr>
<tr>
<td>Older adults</td>
<td>n=5</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>n=4</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>n=4</td>
</tr>
<tr>
<td>Pregnant or postnatal or post-partum women</td>
<td>n=4</td>
</tr>
<tr>
<td>Adults with various chronic diseases</td>
<td>n=3</td>
</tr>
<tr>
<td>Apparently healthy adults</td>
<td>n=2</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td>n=3</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>n=3</td>
</tr>
<tr>
<td>Older adults with depression</td>
<td>n=3</td>
</tr>
<tr>
<td>Renal disease</td>
<td>n=3</td>
</tr>
<tr>
<td>Stroke</td>
<td>n=2</td>
</tr>
<tr>
<td>Arthritis and rheumatic diseases</td>
<td>n=2</td>
</tr>
<tr>
<td>Cognitive impairment</td>
<td>n=2</td>
</tr>
<tr>
<td>HIV/AIDS</td>
<td>n=2</td>
</tr>
<tr>
<td>Post-traumatic stress disorder</td>
<td>n=2</td>
</tr>
<tr>
<td>Various mental health disorders</td>
<td>n=2</td>
</tr>
<tr>
<td>Fatigue</td>
<td>n=1</td>
</tr>
<tr>
<td>Knee osteoarthritis</td>
<td>n=1</td>
</tr>
<tr>
<td>Neurological disorders</td>
<td>n=1</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>n=1</td>
</tr>
<tr>
<td>Substance use disorder</td>
<td>n=1</td>
</tr>
</tbody>
</table>

### Figure 1 Results of meta-analyses that assessed symptoms of depression using standardised mean differences (negative values represent a reduction in symptoms).
Subgroup analyses: clinical status

Depression

Seventeen reviews provided data on patients with cancer, anxiety, and 16 on people with depression or depressive symptoms. PA was effective in reducing depressive symptoms across all conditions (median SMD range: –0.85 (kidney disease), –0.16 (cardiovascular disease)). The largest effects were found in kidney disease, HIV, chronic obstructive pulmonary disease, generally healthy adults and individuals diagnosed with depression (table 2).

Anxiety

PA was generally effective for reducing anxiety across disease conditions, with median SMDs ranging from –1.23 (HIV) to –0.16 (multiple sclerosis). However, the evidence base was limited except for cancer and anxiety disorders (table 3).

Exercise mode

Depression

Eighteen reviews provided analyses by exercise mode (310 component RCTs, >14,496 participants, online supplemental eFigure 6). All modes were effective, and median effect sizes (SMDs) were similar across modes: –0.64 (IQR=–0.86 to –0.40) for strength-based interventions (nine reviews); –0.47 (IQR=–0.64 to –0.29) for mixed-mode interventions (12 reviews); –0.46 (IQR=–0.77 to –0.33) for stretching, yoga and other mind–body modalities (11 reviews) and –0.45 (IQR=–0.79 to –0.37) for aerobic exercise (15 reviews).

Anxiety

Fifteen reviews reported analyses by exercise mode (115 component RCTs, >5451 participants, online supplemental eFigure 7). All modes were effective, with median SMDs of –0.23 (IQR=–0.37 to –0.08) for strength-based interventions (two reviews); –0.35 (IQR=–0.86 to –0.23) for mixed modes (four reviews); –0.42 (IQR=–0.78 to –0.16) for stretching, yoga, and other mind–body modalities (seven reviews) and –0.29 (IQR=–0.54, –0.16) for aerobic exercise (six reviews).

Exercise intensity

Depression

Five reviews reported analyses by exercise intensity (63 component RCTs, >2776 participants, online supplemental eFigure 8). Low, moderate and high-intensity exercise interventions had a median SMD of –0.22 (IQR=–0.50 to –0.12), –0.56 (IQR=–1.03 to –0.33) and –0.70 (IQR=–1.25 to –0.24), respectively.

Anxiety

Two reviews reported analyses by exercise intensity (23 component RCTs, online supplemental eFigure 9). All intensities were effective. The single review for low-intensity exercise had a median SMD of –0.26; the one for moderate-intensity exercise –0.47, and the two for high-intensity exercise –0.44 (IQR=–0.49 to –0.13).

Intervention duration

Depression

Twelve reviews reported analyses by intervention duration (166 component RCTs, 15,669 participants, online supplemental eFigure 10). All durations were effective, but effectiveness declined as intervention duration increased. The median SMDs for short (≤12 weeks, 12 reviews), medium (12–23 weeks, 11 reviews) and long duration (≥24 weeks, 4 reviews) interventions were –0.84 (IQR=–1.50 to –0.48), –0.46 (IQR=–0.53 to –0.25) and –0.28 (IQR=–1.15 to –0.17), respectively.

Anxiety

Four reviews reported analyses by intervention duration (38 component RCTs, 2325 participants, online supplemental eFigure 11). Median SMDs for short (12 weeks) and medium-duration (12–23 weeks) interventions were –0.55 (IQR=–0.83 to –0.27) and –0.47 (IQR=–0.72 to –0.08), respectively. The single review reporting on longer interventions (≥24 weeks) reported a median SMD of –0.15.

Weekly duration

Depression

Four reviews presented analyses by weekly session duration (68 component RCTs, >5016 participants, online supplemental eFigure 12). The median SMD for ≤150 min/week and >150 min/week was –0.58 (IQR=–0.77 to –0.30) and –0.29 (IQR=–0.40 to –0.07), respectively.

Anxiety

One review provided analyses by weekly session duration (17 component RCTs, online supplemental eFigure 13). The median SMDs for <150 min/week and ≥150 min/week were –1.23 and –0.99, respectively.

Session frequency

Depression

Three reviews reported analyses by session frequency. High-frequency (5–7 sessions per week), moderate-frequency (4–5 per week) and low-frequency (<4 per week) interventions had a median SMD of –0.76 (IQR=–1.20 to –0.32), –1.12 (–1.39 to –0.85) and –0.47 (IQR=–0.59 to –0.35), respectively (online supplemental eFigure 14).

Anxiety
One review\(^7^8\) compared session frequency, with SMDs of –0.50, –0.96 and –0.52 for 2–3, 4–5 and 6–7 session per week, respectively (online supplemental eFigure 13).

Session duration
Depression
Three reviews\(^4^2\)\(^\text{--}\)\(^7^8\) presented analyses on session duration (online supplemental content 17). Long (≥60 min, SMD=–0.57, IQR =–0.85 to –0.35) and medium (30–60 min, SMD=–0.60, IQR =–0.78 to –0.41) session durations had similar benefits. The sole study of short sessions (<30 min) had a SMD of 0.01 (online supplemental eFigure 15).

Anxiety
One review\(^7^8\) reported on the effects of session duration (online supplemental content 16). There was no difference between long (SMD=–0.63) and short (SMD=–0.83) sessions (online supplemental eFigure 13).

**DISCUSSION**
This is the first ever study to compile the extensive base of evidence regarding the effects of PA on depression, anxiety and psychological distress. We identified 97 systematic reviews, reporting the findings of 1039 unique RCTs, involving 128 119 participants. Findings suggest that PA interventions are effective in improving symptoms of depression and anxiety. Improvements were observed across all clinical populations, though the...
magnitude of effect varied across different clinical populations. The greatest benefits were seen in people with depression, pregnant and postpartum women, apparently healthy individuals and individuals diagnosed with HIV or kidney disease. All PA modes were effective, and higher intensity exercise was associated with greater improvements for depression and anxiety. Longer duration interventions had smaller effects compared with short and mid-duration, though the longest duration interventions still had positive effects.

PA was effective at reducing depression and anxiety across all clinical conditions, though the magnitude of the benefit varied between clinical groups. The larger effect sizes observed in clinical populations may reflect that these populations experience above-average symptoms of depression and anxiety and have low PA levels, and, therefore, have a greater scope for improvement compared with non-clinical populations.17

All PA modes were beneficial, including aerobic, resistance, mixed-mode exercise and yoga. It is likely that the beneficial effects of PA on depression and anxiety are due to a combination of various psychological, neurophysiological and social mechanisms.87 Different modes of PA stimulate different physiological88 and psychosocial effects,88–90 and this was supported by our findings (eg, resistance exercise had the largest effects on depression, while Yoga and other mind–body exercises were most effective for reducing anxiety). Furthermore, our findings showed that moderate-intensity and high-intensity PA modes were more effective than lower intensities. PA improves depression through various neuromolecular mechanisms including increased expression of neurotrophic factors, increased availability of serotonin and norepinephrine, regulation of hypothalamic–pituitary–adrenal axis activity and reduced systemic inflammation.9192 Therefore, low-intensity PA may be insufficient for stimulating the neurological and hormonal changes that are associated with larger improvements in depression and anxiety.87 Overall, our findings add further support to public health guidelines, which recommend multimodal, moderate and vigorous intensity PA.

Our findings that longer duration interventions were less effective than shorter interventions may seem counter intuitive. It is possible that this finding reflects a decline in adherence with longer interventions. Furthermore, due to a lack of blinding of participants in PA trials, participants may have expected to have improved symptoms. It is possible that after experiencing short-term improvements in depression or anxiety, the expectancy effect may diminish over longer periods of time. An alternative explanation is that the longer interventions might not provide sufficient progression of PA dose, leading to a reduction in their effectiveness. Furthermore, it was somewhat surprising that smaller weekly duration interventions demonstrated larger effects than higher weekly duration. This is the opposite to the dose–benefit relationship observed for exercise and physical health outcomes.93 It is possible that shorter duration interventions are easier for participants to comply with, whereas longer weekly duration interventions are more burdensome and that may be impacting the psychological benefits. It is a useful message that interventions do not need to provide high doses of PA for improvements in depression.

The key strength of this study was that it is the first umbrella review to evaluate the effects of all types of PA on depression, anxiety and psychological distress in all adult populations. We included only the highest level of evidence: meta-analyses of RCTs and applied stringent criteria regarding the design of the component RCTs to ensure that effects could be confidently attributed to PA rather than other intervention components. Additionally, there was only slight overlap in the component RCTs, increasing our confidence in the findings.

A limitation of the review is that most evidence focused on mild-to-moderate depression, with fewer reviews addressing anxiety and psychological distress, preventing us from reaching...
firm conclusions in the subgroup analyses for these outcomes. Furthermore, most (n=77) of the included reviews were rated as ‘critically low’, based on the AMSTAR-2 quality rating.

Clinical implications

PA is effective for managing symptoms of depression and anxiety across numerous populations, including the general population, people with mental illnesses and various other clinical populations. While the benefit of exercise for depression and anxiety is generally recognised, it is often overlooked in the management of these conditions. Furthermore, many people with depression and anxiety have comorbidities, and PA is beneficial for their mental health and disease management. This underscores the need for PA to be a mainstay approach for managing depression and anxiety.

All modes of PA are effective, with moderate-to-high intensities more effective than low intensity. Larger benefits are achieved from shorter interventions, which has health service delivery cost implications—suggesting that benefits can be obtained following short-term interventions, and intensive long-term interventions are not necessarily required to achieve therapeutic benefit. The effect size reductions in symptoms of depression (−0.43) and anxiety (−0.42) are comparable to or slightly greater than the effects observed for psychotherapy and pharmacotherapy (SMD range = −0.22 to −0.37). Future research to understand the relative effectiveness of PA compared with (and in combination with) other treatments is needed to confirm these findings.

In conclusion, PA is effective for improving depression and anxiety across a very wide range of populations. All PA modes are effective, and higher intensity is associated with greater benefit. The findings from this umbrella review underscore the need for PA, including structured exercise interventions, as a mainstay approach for managing depression and anxiety.

What is already known

⇒ Previous research trials suggest that physical activity may have similar effects to psychotherapy and pharmacotherapy for patients with depression, anxiety or psychological distress.
⇒ Studies have evaluated different forms of physical activity, in varying dosages, in different population subgroups, and using different comparator groups, making it difficult for clinicians to understand the body of evidence for physical activity in the management of mental health disorders.

What are the new findings

⇒ Results showed that physical activity is effective for reducing mild-to-moderate symptoms of depression, anxiety and psychological distress (median effect size range = −0.42 to −0.60), compared with usual care across all populations.
⇒ Our findings underscore the important role of physical activity in the management of mild-to-moderate symptoms of depression, anxiety and psychological distress.

REFERENCES

Womens Health (Larchmt) 2020;9:1325.


Review


