Does aerobic exercise reduce postpartum depressive symptoms? A systematic review and meta-analysis

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This paper was accepted for publication in the journal British Journal of General Practice and the definitive published version is available at https://doi.org/10.3399/bjgp17X692525

Metadata Record: https://dspace.lboro.ac.uk/2134/32201

Version: Accepted for publication

Publisher: Royal College of General Practitioners © British Journal of General Practice

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Does aerobic exercise reduce postpartum depressive symptoms?

A systematic review and meta-analysis

Running Title: Systematic review of exercise for postpartum depressive symptoms

Ruth Victoria Pritchett, Amanda J Daley, Kate Jolly


Kate Jolly: Institute of Applied Health Research, Public Health Building, The University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK. Funded by The Collaboration for Leadership in Applied Health Research and Care West Midlands.

Correspondence and reprint requests to Kate Jolly: Phone 0121 414 7552, Email: c.b.jolly@bham.ac.uk

Word count: 2719
ABSTRACT

Background
There is currently no specific guidance on the role of exercise in managing postpartum depression in the UK and USA and international guidance is inconsistent.

Aim
To assess the effectiveness of aerobic exercise on postpartum depressive symptoms.

Design and setting
Systematic review and meta-analysis including randomised controlled trials (RCTs) of interventions to increase exercise in the postpartum period. There was no restriction to study site or setting.

Method
The databases MEDLINE, EMBASE, Cochrane Library, PsychINFO, SportDiscus, http://clinicaltrials.gov and the World Health Organisation International Clinical Trials Registry Platform were searched. Titles and abstracts, then full text articles were screened against inclusion criteria: RCTs measuring depressive symptoms in mothers ≤ 1 year postpartum; interventions designed to increase aerobic exercise compared with usual care or other comparators. Included studies were assessed using the Cochrane Collaboration’s risk of bias tool. Meta-analysis was conducted. Pre-planned subgroup analyses explored heterogeneity.

Results
Thirteen RCTs were included, with 1734 eligible participants. Exercise significantly reduced depressive symptoms when all trials were combined (standardized mean difference (95% confidence interval) -0.44, (-0.75 to -0.12). Exploration of heterogeneity did not find significant differences in effect size between women with possible depression and in general postpartum populations; exercise-only and exercise with co-interventions; and group exercise and exercise counselling.
Conclusion

This systematic review provides support for the effectiveness of exercise in reducing postpartum depressive symptoms. Group exercise, participant-chosen exercise and exercise with co-interventions may all be effective interventions. These results should be interpreted with caution due to substantial heterogeneity and risk of bias.

Key Words:
Depression, Postpartum, Postnatal, Exercise, Systematic review

How this fits in

UK clinical guidance recommends psychological therapy and antidepressants for postnatal depression, but women can be reluctant to take antidepressants postnatally and the availability of psychological therapies is often limited. Exercise is effective in treating depression in the general adult population, however, there are particular challenges in the postnatal period which may impact on the feasibility and effectiveness of exercise in reducing postnatal depressive symptoms. This systematic review found that exercise is an effective intervention for reducing postnatal depressive symptoms. Encouraging women to take aerobic exercise in the postnatal period may have psychological and other health benefits.
Introduction

Postpartum depression is a global mental health issue, annually affecting 13 million women worldwide (1-3). Postpartum depression can lead to thoughts of self-harm and/or harm of the child (4) and negatively affect children’s development (5, 6). There can be reluctance by women to take antidepressants postnatally, particularly if they are breastfeeding (7, 8). Cognitive behavioral therapy can be effective but waiting lists are often long in the United Kingdom (UK) (9).

Exercise can have a moderate effect in reducing general adult depression (standardized mean difference -0.62, 95% confidence interval -0.81 to -0.42) (10). Practitioner led group exercise is recommended by the UK National Institute for Health and Care Excellence (NICE) for persistent subthreshold depressive symptoms and mild to moderate depression (11). However, postpartum women have particular physical, practical and psychological challenges to exercising (12, 13). A 2009 review provided some support for exercise as an adjunctive treatment for postpartum depressive symptoms (14).

Despite its challenges, exercise provides another intervention option when many mothers are reluctant to seek treatment and/or take antidepressants (8). Exercise also offers improved fitness and weight loss and opportunities for social interaction at a time when women experience decreased exercise and weight retention (15-17).

Aim and objectives

To undertake a meta-analysis of randomized controlled trials (RCTs) investigating the effect of aerobic exercise interventions, compared to usual care, on depressive symptoms in women up to one year postpartum.

To explore the effect of exercise on depressive symptoms in three subgroup analyses: (i) women with possible depression and general postpartum populations; (ii) interventions based only on exercise and those with co-interventions; and (iii) interventions providing structured group exercise and those supporting participant choice of exercise.
Methods

The review was registered on PROSPERO: CRD42016047656.

Data sources and searches

We followed the Cochrane Collaboration’s guidance on reviews of interventions (18) and PRISMA guidelines (19). We searched the following bibliographic databases electronically for eligible trials: MEDLINE, EMBASE, the Cochrane Library, PsychINFO, SportDiscus and Clinical Trials.gov. The World Health Organisation (WHO) International Clinical Trials Registry Platform was searched for trials in progress. We conducted searches for grey literature in Open Grey (20) and OAIster Worldcat (21).

Searches were conducted on a range of psychological outcomes, only the results for depression are presented. We searched for the following terms as text words and MeSH terms where applicable: exercise, physical activity, postpartum, postnatal, mother, birth, perinatal, depression, anxiety, self-efficacy, quality of life, mother and infant bonding and child development (for Medline search strategy see Appendix 1). Searches were unrestricted by date or language and undertaken up to September 2016. We searched bibliographies of eligible studies and reviews for additional trials and contacted authors if required. Two researchers independently reviewed titles, abstracts and full text articles. Eligibility discrepancies were discussed with a third reviewer.

Study Selection

Inclusion criteria were: (i) populations of mothers less than one year postpartum; (ii) interventions designed to increase aerobic exercise (activity causing increased heart rate, respiratory rate and sweating), including those with co-interventions such as social support or nutrition elements; (iii) comparator groups receiving no care or any form of usual care; (iv) depressive symptoms measured by questionnaire or diagnostic interview; (v) RCTs.

Exclusion criteria: Trials comparing two types of exercise.

Data extraction and quality assessment
Data describing the population, intervention and findings were extracted independently by two researchers using standardized, piloted, extraction spreadsheets. Reviewers were not blinded to authors, institution, or journal. Authors were contacted if clarification was required.

Two researchers independently applied the Cochrane Collaboration’s risk of bias tool (18) in relation to sequence generation; allocation concealment; blinding of outcome assessors; incomplete outcome data; selective outcome reporting and other sources of bias (18). Blinding of personnel and participants to group allocation was not assessed due to the impracticalities of this in exercise trials. Each study was allocated an overall risk of bias, with a low risk of bias being assigned if the risk of bias was deemed to be low in all sub domains; an unclear risk of bias if the risk was considered unclear in one or more subdomains and a high risk of bias if the risk was considered high in one or more subdomains (18). Publication bias was assessed by visually inspecting a funnel plot of the standardized mean differences (SMD) for depressive symptoms.

Data synthesis and analysis

Populations indicating possible depression on screening questionnaires or in the clinical judgement of a health professional are referred to as ‘depressed’ postpartum populations.

We conducted meta-analyses and meta-regressions using Review Manager 5.3 (22). We calculated an SMD (18) for continuous outcomes of depressive symptoms. If more than one measure for depression was reported, we used the primary continuous outcome measure. The final point of follow up was used as we wished to report long term outcomes. We calculated a weighted mean difference (WMD) for all trials using the Edinburgh Postnatal Depression Scale (EPDS) to assess depressive symptoms (23). Where the standard deviation of the difference in score was not reported, we calculated this. A correlation of 0.6 was used in the assumptions to estimate standard deviations of differences, based on a previous sample of postpartum women (24). We used a random effects model throughout
because of known clinical heterogeneity between populations and methodological
heterogeneity between interventions.

We calculated an initial pooled standardized mean difference with 95% CI, with all trials
included. Statistical significance was defined as having a 95% CI that did not include zero. We
explored clinical heterogeneity (by qualitatively comparing characteristics between trials) and
statistical heterogeneity using chi-squared tests of heterogeneity and the $I^2$ statistic.

We performed three a priori subgroup analyses, investigating whether the effectiveness of
exercise in reducing postpartum depressive symptoms varied in relation to (i) the population
of women (women with possible depression or general postpartum populations); (ii) the
presence of co-interventions (exercise–only or exercise with co-interventions such as diet or
social support) and (iii) the context of exercise (exercise groups or participant choice,
participant choice often consisted of exercise counselling interventions in which participants
typically elected to exercise alone, often by walking). For each hypothesis we categorised
the trials into subgroups and used a random effects meta-regression model to determine the
effects of exercise in each subgroup and the significance of differences between subgroups.

Results

Trial selection

Of the 9165 records identified after the removal of duplicates, 9043 were excluded after
reading titles and abstracts. A further 109 full text records were excluded based on the
eligibility criteria. Thirteen trials were included in the meta-analysis (Figure 1).

Trial characteristics

All 13 RCTs had been peer reviewed and published as journal articles (25-36) or abstracts
(37). Four trials were conducted in the USA (28, 30, 32, 34); three in England (31, 36, 37);
two in Australia (25, 29); one in Canada (26); one in India (33); one in Japan (35) and one in
Taiwan (27).
The total population of the combined studies was 1734 eligible participants, with follow up data provided by 1307 participants who were included in the primary meta-analysis. Seven trials recruited participants with possible depression (25-28, 31, 36, 37), of which three required participants to score above a threshold on the EPDS questionnaire (from ≥ 10 to >12) (29,30,35) and two required a diagnosis of depression (ICD-10 (36) or DSM-IV criteria (37)). Despite not reporting baseline depression threshold criteria, two trials reported mean baseline depression scores indicating depressed populations (EPDS scores of 18.9 - 19.8 (28), BDI 15.8 - 16.9 (27)). These trials were therefore considered to have ‘depressed’ postpartum populations. Six trials recruited general postpartum populations (29, 30, 32-35) (Table 1).

**Intervention characteristics**

All trials included interventions designed to increase exercise levels. In addition, two trial interventions also aimed to improve diet and encourage a healthy lifestyle (27, 30); two also provided social support (25, 34) and one also provided education on postpartum issues (29). Six trials had group exercise interventions (25, 29, 33-35, 37); seven trials had interventions in which exercise counselling was provided and the participant was free to choose their own form of exercise (26-28, 30-32, 36) (Table 1). The interventions ranged from 4 weeks duration (33,35) to 6 months (27,32,36). Most aimed to achieve 30 minutes of moderate activity 3-5 times weekly (25,26,28,30-32,35-37), although some consisted of once weekly group exercise (29,33). Eight of the studies reported adherence to the intended intervention (25,26,28,31,32,35-37).

Risk of bias in included studies and publication bias

Six trials were considered at unclear risk (25, 26, 31-33, 37) and seven at high risk of bias (27-30, 34-36). Principal factors introducing a risk of bias were a lack of intention-to-treat analyses (some studies excluded non-adherent participants), a lack of clarity on selective outcome reporting, a lack of robust sequence generation and concealment of randomisation
procedures and unclear blinding of those conducting outcome assessments and analyses (Appendix 2).

We visually inspected the funnel plot for the main analysis including all trials. There appeared to be a lack of smaller studies with results indicating an increase in depression in exercise intervention groups compared to comparator groups. This may be an indication of publication bias (Appendix 3).

**Provision of additional data**

Authors of four trials provided additional data for this review: DaCosta et al (26), Surkan et al (30) Lewis et al (32) and Boath et al (37).

**Data analysis**

**Overall effect of exercise on depression scores**

Exercise interventions significantly reduced depressive symptoms (SMD -0.44, 95% CI -0.75 to -0.12, n=1307, I²: 85%, 13 trials) (Figure 2) (25-37). The WMD was -1.54 EPDS units, 95% CI -2.97 to -0.12, n=652, I²: 87% (25, 26, 28, 29, 31, 33-37).

**Population**

Exercise interventions had a significant effect in reducing depressive symptoms in 'depressed' postpartum populations (SMD -0.32, 95% CI -0.63 to -0.00), I²: 55%, (25-28, 31, 36, 37) (Table 2, Appendix 4) and in general postpartum populations (-0.57, 95% CI -1.12 to -0.02, I²: 92%) (29, 30, 32-35) (Table 2, Appendix Figure 3). The effect of exercise interventions in the 'depressed' and general postpartum populations was not significantly different (test for subgroup differences X²= 0.62, P= 0.43, I²: 0%) (Appendix 4).

**Intervention type**

Exercise-only interventions had a non-significant effect in reducing depressive symptoms (SMD -0.56, 95% CI -1.13 to 0.01, I²: 89%) (26, 28, 31-33, 35-37). Exercise with co-interventions had a significant effect on reducing depressive symptoms (-0.35, 95% CI -0.66 to -0.04, I²: 72%) (25, 27, 29, 30, 34). The effect of exercise-only interventions and exercise...
co-interventions on depressive symptoms was not significantly different (test for subgroup differences: $X^2 = 0.41$, $P = 0.52$, $I^2: 0\%$) (Table 2, Appendix 5).

**Exercise context**

Group exercise interventions had a significant effect in reducing depressive symptoms (standardized mean difference $-1.10$, 95% CI $-1.99$ to $-0.21$, $I^2: 93\%$) (25, 29, 33-35, 37). Participant choice interventions such as exercise counselling with personal choice of exercise (often exercise alone) had a significant effect in reducing depressive symptoms ($-0.20$, 95% CI $-0.33$ to $-0.06$, $I^2: 0\%$) (26-28, 30-32, 36). The effects of group exercise and participant choice exercise on depressive symptoms were not significantly different (test for subgroup differences: $X^2 = 3.89$, $P = 0.05$, $I^2: 74\%$) (Table 2, Figure 6).

**Discussion**

**Summary**

There is no specific guidance on the role of exercise in the management of perinatal mental health in the UK or USA (7, 38), this systematic review provides support for exercise as an effective treatment for reducing postpartum depressive symptoms, whether or not women meet robust criteria for PND. Additionally, we have been able to explore the characteristics of exercise interventions which are most likely to be effective for the postpartum population.

**Strengths and limitations**

Caution should be taken when interpreting our results, as a substantial level of heterogeneity was present. Heterogeneity was present in the design of the trials, including exercise-only interventions and those also promoting a healthy diet or social support amongst peers. Exercise contexts also varied between structured, group-based exercise and tailored exercise counselling with participant choice of exercise. A random effects model was used in analyses to account for this variation and potential causes of heterogeneity were explored in subgroup analyses, however, there were insufficient numbers of trials to explore this
heterogeneity in depth. We used the last follow-up point as we were interested in long term outcomes; the median duration of follow-up was 6 months from recruitment (IQR 3,9).

The methodological quality of several of the included trials was low. Exclusion of non-adherent participants, insufficiently robust sequence generation and unclear blinding of outcome assessors resulted in an increased risk of bias. Only two included trials in ‘depressed’ populations recruited women with a diagnosis of depression (ICD-10 or DSM-IV) (36, 37); the remainder used screening questionnaires such as the EPDS (23) to indicate women with possible depression. The EPDS has the advantage of being relatively short and simple to complete, but has been reported to have considerable heterogeneity in sensitivity and specificity for detecting PND across different settings and populations (39).

There was a lack of assessment of exercise duration and intensity in included studies; of those that did provide this information, only two included an objective measure of exercise (34, 36). Exercise intensity can be difficult to determine accurately by self-report. For future research, objective accelerometry would allow for greater accuracy and more meaningful comparison between intervention types.

This review has several strengths. Recommended methods were followed; searching was systematic and not limited by language of publication. We restricted the review to RCTs to reduce the potential for confounding. We explored the effectiveness of exercise in different populations; intervention types and exercise contexts, which has not been previously attempted.

Comparison with existing literature

The evidence in this field has increased substantially since the review by Daley et al. in 2009 (14). Although the primary effect size in this review is smaller than that reported by Daley et
al. (SMD -0.44, 95%CI -0.75 to -0.12 rather than -0.81, 95%CI -1.53 to -0.10 (14)), the CI surrounding the effect is narrower here, indicating more precision.

In this review, the effect of exercise in depressed postpartum populations was smaller than that reported in a recent review of exercise in general adult depressed populations (SMD -0.62, 95%CI -0.81 to -0.42) (10), indicating that there may be differences in the way general adult populations and postpartum populations respond to exercise or differing abilities to adhere to an exercise regime. Women in the postnatal period have the time constraints of new parenthood and barriers to exercise such as fatigue. These factors could be explored by qualitative research with postnatal women.

The effect of exercise on postpartum depressive symptoms reported in this review compared favourably with that reported in a review of low intensity psychological interventions (such as online CBT and self-help literature) versus usual care for depression in the general population (SMD -0.42, 95%CI -0.55 to -0.29) (40). Despite some temporary changes in the composition of breastmilk following maximal exercise, the literature does not suggest any detrimental effects of moderate levels of aerobic exercise (41), as evaluated by the trials in this review. Such interventions are often amongst the first treatments offered to individuals presenting with depression (11). The effect of exercise on postpartum depressive symptoms we report is also comparable to those reported for antidepressants in depressed adult populations; (SMD -0.49, 95%CI -0.67 to -0.32 for tricyclic antidepressants and -0.24, 95%CI -0.35 to -0.12 for selective serotonin reuptake inhibitors compared to placebo) (42).

Exercise interventions also offer the additional health benefits of improved physical fitness and weight loss, outcomes particularly relevant to postpartum women, who often experience decreased exercise levels and excess weight (15-17). The finding that both group exercise and participant choice of exercise (often solitary exercise) may be effective in reducing postpartum depressive symptoms is also noteworthy. Current guidance from NICE on the treatment of postpartum mental illness refers clinicians to exercise advice for the general
adult population recommending only group exercise (11), which may not be the most feasible or acceptable option for women after having a baby (43).

Implications for research and practice

National and international guidance on the use of exercise to reduce depressive symptoms in the postpartum period is lacking. Whilst acknowledging considerable uncertainty about our findings due to substantial heterogeneity, given the high prevalence of postpartum depression and the potential for exercise to be a low cost, freely available intervention, aerobic exercise should be considered as a management option for postpartum women with depressive symptoms and as a potentially prevention measure more generally in postpartum women.

Funding Source

This paper presents independent research funded by the University of Birmingham, The National Institute for Health Research (NIHR) School for Primary Care Research and The Collaboration for Leadership in Applied Health Research and Care West Midlands. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Ethical approval

Not applicable

Competing interests

All authors declare no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work. Dr Pritchett (née Blamey) reports I am an author of one of the trials in this review, Dr Daley
reports I am an author of two of the trials included in this review, Dr Jolly reports I am an author of one of the trials included in this review.

Acknowledgements

Authors RP, KJ and AD conceived the study. RP designed the search strategy. All authors selected the papers for inclusion and abstracted the data. RP undertook the meta-analysis under the guidance of AD. RP wrote the first draft of the paper. All authors had critical input into the writing and approved the final version. No other persons contributed significantly to this work.

References


Figure 1: PRISMA flow chart

Number of records identified through database searching: 9300

Number of additional records identified from bibliographies of eligible papers: 38

Number of additional records identified while searching for full text versions of papers: 1

Number of additional records identified by communication with the authors: 1

Number of records after duplicates removed: 9165

Number of records screened: 9165

Number of records excluded: 9043
(Abstracts not relevant to this review)

Number of full-text articles excluded: 109
Reasons for exclusion:
63 (58%): Not an RCT
11 (10%): Intervention not aerobic exercise
2 (2%): Intervention and control were two different types of exercise, no comparator without exercise present
21 (19%): Depressive symptoms not measured
12 (11%): Population not women less than one year postpartum

Number of full-text articles assessed for eligibility: 122

Number of trials included in review: 13
Figure 2: Meta-analysis of the effect of exercise on depressive symptoms

(Standardized mean difference)

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Exercise Mean</th>
<th>Exercise SD</th>
<th>Exercise Total</th>
<th>Comparator Mean</th>
<th>Comparator SD</th>
<th>Comparator Total</th>
<th>Std. Mean Difference</th>
<th>IV, Random, 95% CI</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armstrong 2003</td>
<td>-12.8</td>
<td>3.76</td>
<td>10</td>
<td>-3.7</td>
<td>6.13</td>
<td>10</td>
<td>-1.71</td>
<td>[-2.77, -0.88]</td>
<td>2003</td>
</tr>
<tr>
<td>Robichaud 2008</td>
<td>-1.68</td>
<td>3.74</td>
<td>25</td>
<td>-0.49</td>
<td>3.11</td>
<td>23</td>
<td>-0.34</td>
<td>[-0.91, 0.23]</td>
<td>2008</td>
</tr>
<tr>
<td>Daley 2008</td>
<td>-4.8</td>
<td>4.81</td>
<td>16</td>
<td>-4.5</td>
<td>4.8</td>
<td>15</td>
<td>-0.08</td>
<td>[-0.77, 0.64]</td>
<td>2008</td>
</tr>
<tr>
<td>De Costa 2009</td>
<td>-6.39</td>
<td>3.98</td>
<td>46</td>
<td>-5.2</td>
<td>4.04</td>
<td>42</td>
<td>0.65</td>
<td>[0.37, 0.93]</td>
<td>2009</td>
</tr>
<tr>
<td>Norman 2010</td>
<td>-3.27</td>
<td>5.17</td>
<td>82</td>
<td>-0.21</td>
<td>4.94</td>
<td>73</td>
<td>-0.60</td>
<td>[-0.95, -0.25]</td>
<td>2010</td>
</tr>
<tr>
<td>Huang 2011</td>
<td>3.01</td>
<td>5.87</td>
<td>84</td>
<td>4.5</td>
<td>7.02</td>
<td>84</td>
<td>-0.23</td>
<td>[-0.56, 0.12]</td>
<td>2011</td>
</tr>
<tr>
<td>Surkan 2012</td>
<td>-1.3</td>
<td>9.86</td>
<td>203</td>
<td>1.33</td>
<td>10.9</td>
<td>200</td>
<td>-0.24</td>
<td>[-0.44, -0.05]</td>
<td>2012</td>
</tr>
<tr>
<td>Hanada 2013</td>
<td>-0.6</td>
<td>3.67</td>
<td>48</td>
<td>-1.8</td>
<td>3.24</td>
<td>47</td>
<td>0.37</td>
<td>[-0.03, 0.77]</td>
<td>2013</td>
</tr>
<tr>
<td>Thiruppathi 2014</td>
<td>-3.04</td>
<td>0.64</td>
<td>20</td>
<td>-0.24</td>
<td>0.52</td>
<td>21</td>
<td>-4.60</td>
<td>[-5.88, -3.32]</td>
<td>2014</td>
</tr>
<tr>
<td>Lewis 2014</td>
<td>-2.08</td>
<td>2.75</td>
<td>61</td>
<td>-1.27</td>
<td>3.97</td>
<td>63</td>
<td>-0.24</td>
<td>[-0.59, 0.12]</td>
<td>2014</td>
</tr>
<tr>
<td>Keller 2014</td>
<td>-1.21</td>
<td>4.9</td>
<td>39</td>
<td>-1.71</td>
<td>4.04</td>
<td>54</td>
<td>0.11</td>
<td>[-0.30, 0.52]</td>
<td>2014</td>
</tr>
<tr>
<td>Daley 2015</td>
<td>-5.23</td>
<td>4.24</td>
<td>41</td>
<td>-4.95</td>
<td>4.18</td>
<td>38</td>
<td>0.52</td>
<td>[-0.52, 0.08]</td>
<td>2015</td>
</tr>
<tr>
<td>Boath 2015</td>
<td>-8.9</td>
<td>5.52</td>
<td>11</td>
<td>-3.2</td>
<td>4.68</td>
<td>11</td>
<td>1.07</td>
<td>[-1.98, -0.17]</td>
<td>2015</td>
</tr>
</tbody>
</table>

Total (95% CI) 646 661 100.0% -0.44 [-0.75, -0.12]

Heterogeneity: $I^2 = 0.26$, $Chi^2 = 79.17$, df = 12 ($P < 0.000001$); $P = 95$
Test for overall effect: $Z = 2.72$ ($P = 0.006$).
<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Population</th>
<th>N</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Outcomes</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armstrong</td>
<td>Queensland</td>
<td>Depressed population</td>
<td>20</td>
<td>Group exercise programme (pram walks) with social co-intervention 12 weeks</td>
<td>Phone support at 6 weeks post randomization</td>
<td>EPDS and DASS 6 &amp; 12 weeks</td>
<td>Intervention group depression improved significantly compared to control</td>
</tr>
<tr>
<td>Daley</td>
<td>Birmingham</td>
<td>SEDENTARY POPULATION</td>
<td>38</td>
<td>Exercise counselling 12 weeks</td>
<td>Usual care Exercise consult at end of trial</td>
<td>EPDS 24 weeks</td>
<td>No significant difference in EPDS between groups</td>
</tr>
<tr>
<td>Robichaud</td>
<td>Tennessee</td>
<td>Depressed population</td>
<td>48</td>
<td>Exercise programme (home-based walking) 6 weeks</td>
<td>Usual activity Intervention after trial</td>
<td>EPDS 8 weeks</td>
<td>No significant difference in EPDS between groups</td>
</tr>
<tr>
<td>Da Costa</td>
<td>Montreal</td>
<td>Depressed population</td>
<td>88</td>
<td>Exercise counselling 12 weeks</td>
<td>Usual care</td>
<td>EPDS and 17 item HAM-D 3 &amp; 6 months</td>
<td>No significant change in EPDS from baseline to follow up.</td>
</tr>
<tr>
<td>Norman</td>
<td>Melbourne</td>
<td>General postpartum</td>
<td>161</td>
<td>Group exercise programme with co-intervention (weekly educational session on postpartum issues) 8 weeks</td>
<td>Education on topics unrelated to diet or physical activity Booklet of exercises List of gyms and community facilities</td>
<td>EPDS 8 &amp; 12 weeks</td>
<td>Significant decrease in EPDS in intervention compared to comparator</td>
</tr>
<tr>
<td>Huang</td>
<td>Northern</td>
<td>Depressed population</td>
<td>160</td>
<td>Exercise counselling with dietary co-intervention Postpartum group: 6 months</td>
<td>‘Standard care’ Discussions with nurses and leaflets on unrelated topics</td>
<td>BDI 6 months</td>
<td>Significant increase in BDI in intervention and comparator groups</td>
</tr>
<tr>
<td>Surkan</td>
<td>Northeast</td>
<td>General postpartum</td>
<td>679</td>
<td>Exercise counselling with dietary co-intervention</td>
<td>Usual WIC care: nutrition</td>
<td>CES-D Mean follow</td>
<td>Modest decrease in CES-D in intervention</td>
</tr>
<tr>
<td>Study</td>
<td>Location</td>
<td>Population</td>
<td>Sample Size</td>
<td>Intervention</td>
<td>Outcomes</td>
<td>Comparison</td>
<td>Notes</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>------------</td>
<td>-------------</td>
<td>--------------</td>
<td>----------</td>
<td>------------</td>
<td>-------</td>
</tr>
<tr>
<td>Haruna 2013 (35)</td>
<td>Tokyo Japan</td>
<td>General postpartum population</td>
<td>110</td>
<td>Group exercise programme 4 weeks</td>
<td>EPDS 2 &amp; 4 months</td>
<td>Compared to control. Difference was significant in fully adjusted analysis</td>
<td></td>
</tr>
<tr>
<td>Thiruppathi 2014 (33)</td>
<td>Nellore India</td>
<td>General postpartum population</td>
<td>45</td>
<td>Group exercise programme 4 weeks</td>
<td>EPDS 12 weeks</td>
<td>Significant decrease in intervention group from baseline to follow up. No significant change in comparator group</td>
<td></td>
</tr>
<tr>
<td>Keller 2014 (34)</td>
<td>A south western city, USA</td>
<td>General postpartum population</td>
<td>139</td>
<td>Exercise counselling with group exercise (walking) and a co-intervention of ‘promotoras’ (peer support) 12 weeks</td>
<td>EPDS 6 and 12 months</td>
<td>No significant difference between groups</td>
<td></td>
</tr>
<tr>
<td>Lewis 2014 (32)</td>
<td>Minnesota USA</td>
<td>General postpartum population</td>
<td>130</td>
<td>Exercise counselling 6 months</td>
<td>SCID-1 at 6 months only, PHQ-9 EPDS at 6 months only</td>
<td>Significant decrease in PHQ-9 and EPDS in intervention compared to control</td>
<td></td>
</tr>
<tr>
<td>Boath 2015 (37)</td>
<td>Stoke-on-Trent, UK</td>
<td>Depressed population</td>
<td>22</td>
<td>Group exercise programme (pram walk or facility-based) 12 weeks</td>
<td>EPDS 3 and 6 months</td>
<td>No significant difference in EPDS between groups</td>
<td></td>
</tr>
<tr>
<td>Daley 2015</td>
<td>Birmingham</td>
<td>Depressed population</td>
<td>94</td>
<td>Exercise counselling</td>
<td>Usual care</td>
<td>EPDS</td>
<td>Decrease in EPDS in</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Time</td>
<td>Intervention</td>
<td>Timing</td>
<td>Comparison</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>-------</td>
<td>--------------</td>
<td>--------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td></td>
<td>6 months</td>
<td>Leaflet on healthy postpartum lifestyle</td>
<td>6 and 12 months</td>
<td>intervention vs comparator, only significant at 6 months not at 12 months</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BDI: The Beck Depression Inventory; CES-D: The Centre for Epidemiologic Studies Depression Scale; Comp: Comparator; DASS: The Depression Anxiety Stress Scale; EPDS: The Edinburgh Postnatal Depression Scale; HAM-D: The Hamilton Rating Scale for Depression; Int: Intervention; PHQ-9: The Patient Health Questionnaire-9; SCID-1: Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I); WIC: Special Supplemental Nutrition Program for Women, Infants and Children
<table>
<thead>
<tr>
<th>Category</th>
<th>Standardized Mean Difference</th>
<th>95% Confidence Interval</th>
<th>P value</th>
<th>I²</th>
<th>Number of trials</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depressed postpartum populations</td>
<td>-0.32</td>
<td>-0.63 to -0.00</td>
<td>0.05</td>
<td>55%</td>
<td>7</td>
<td>416</td>
</tr>
<tr>
<td>General postpartum populations</td>
<td>-0.57</td>
<td>-1.12 to -0.02</td>
<td>0.04</td>
<td>92%</td>
<td>6</td>
<td>891</td>
</tr>
<tr>
<td>Exercise-only interventions</td>
<td>-0.56</td>
<td>-1.13 to 0.01</td>
<td>0.05</td>
<td>89%</td>
<td>8</td>
<td>528</td>
</tr>
<tr>
<td>Exercise co-interventions</td>
<td>-0.35</td>
<td>-0.66 to -0.04</td>
<td>0.03</td>
<td>72%</td>
<td>5</td>
<td>779</td>
</tr>
<tr>
<td>Group exercise interventions</td>
<td>-1.10</td>
<td>-1.99 to -0.21</td>
<td>0.02</td>
<td>93%</td>
<td>6</td>
<td>406</td>
</tr>
<tr>
<td>Participant choice exercise interventions</td>
<td>-0.20</td>
<td>-0.33 to -0.06</td>
<td>0.003</td>
<td>0%</td>
<td>7</td>
<td>901</td>
</tr>
</tbody>
</table>
APPENDIX 1: Medline search strategy

1. Exercise
2. Exercise.tw
3. Physical activity
4. Physical activity.tw
5. Activity
6. Activity.tw
7. 1 OR 2 OR 3 OR 4 OR 5 OR 6
8. Postpartum
9. Postpartum.tw
10. Postnatal
11. Postnatal.tw
12. Mother
13. Mother.tw
14. Birth
15. Birth.tw
16. Perinatal
17. Perinatal.tw
18. 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17
19. 7 AND 18
20. Depression
21. Depression.tw
22. Postnatal depression
23. Postnatal depression.tw
24. Postpartum depression
25. Postpartum depression.tw
26. Anxiety
27. Anxiety.tw
28. Mother and infant bonding
29. Mother and infant bonding. tw
30. Self efficacy
31. Self efficacy.tw
32. Quality of life
33. Quality of life.tw
34. Child development
35. Child development.tw
36. 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR 31 OR 32 OR 33 OR 34 OR 35
37. 19 AND 36
38. Limit 37 to female
39. Limit 38 to human
Remove duplicates from 39
### APPENDIX 2: Risk of bias of included studies

<table>
<thead>
<tr>
<th>Trial</th>
<th>Sequence generation</th>
<th>Allocation concealment</th>
<th>Blinding of outcome assessor/analysis</th>
<th>Incomplete outcome data</th>
<th>Selective outcome reporting</th>
<th>Overall risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armstrong 2003 (25) (Depressed postnatal population)</td>
<td>Unclear: sealed envelope. Unclear if opaque or sequentially numbered</td>
<td>Unclear: Blinding of outcome assessor unclear. Unclear if researchers conducting analysis blinded</td>
<td>Low: All participants included in analysis (n=20).</td>
<td></td>
<td>Unclear: All listed outcomes reported on. Trial not registered on trial registries (proposed outcomes unavailable)</td>
<td>Unclear</td>
</tr>
<tr>
<td>Daley 2008 (31) (Depressed postnatal population)</td>
<td>Low: computer-generated random list from independent statistician</td>
<td>Unclear: No details</td>
<td>Unclear: Outcomes assessed by postal questionnaire. Unclear if researchers conducting analysis blinded</td>
<td>High: Follow up: 31/38 (81.6%) Missing data balanced across arms. No data on characteristics of drop outs. Intention-to-treat analysis was performed but drop outs were excluded.</td>
<td>Low: Listed primary outcomes reported on. All primary outcomes listed on the ISRCTN reported in results (ISRCTN75708176)</td>
<td>High</td>
</tr>
<tr>
<td>Robichaud 2008 (28) (Depressed postnatal population)</td>
<td>Low: Computerised Adaptive Randomisation Programme with controlled baseline characteristics</td>
<td>Unclear: No details</td>
<td>Unclear: Unclear if outcome assessor blinded. Unclear if researchers conducting analysis blinded</td>
<td>High: 48/51 (94.1%) participants included in analysis. However, not intention-to-treat analysis. Three excluded: 1 non-adherent intervention participant, 1 participant in each group who commenced antidepressants.</td>
<td>High: Anxiety as measured by PPQ not reported. Trial not registered on trial registries (proposed outcomes unavailable)</td>
<td>High</td>
</tr>
<tr>
<td>Trial</td>
<td>Sequence generation</td>
<td>Allocation concealment</td>
<td>Blinding of outcome assessor/analysis</td>
<td>Incomplete outcome data</td>
<td>Selective outcome reporting</td>
<td>Overall risk of bias</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>---------------------------------------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Da Costa 2009 (26) (Depressed postnatal population)</td>
<td>Unclear: stratification in blocks of 4-6 based on baseline depression severity. Unclear how sequence generated</td>
<td>Low: Trial personnel blinded at point of allocation</td>
<td>Unclear: One measure of the primary outcome administered by phone by blinded interviewer. One outcome assessed by postal questionnaire. Unclear if researchers conducting analysis blinded</td>
<td>Low: Follow up: 62/88 (70.45%) Missing data balanced across arms. Non completers had similar baseline demographic and clinical characteristics to completers. Analysis intention-to-treat, missing data imputed (n=88).</td>
<td>Unclear: Listed primary outcome measures reported on. Trial registered after its completion NCT00384943</td>
<td>Unclear</td>
</tr>
<tr>
<td>Norman 2010 (29) (General postnatal population)</td>
<td>Low: computer generated random list in blocks of 16, stratified by parity</td>
<td>Low: consecutively numbered, sealed, opaque envelopes</td>
<td>Low: Researcher scoring outcome questionnaires blinded. Unclear if researchers conducting analysis blinded but if data anonymised before imputing it is likely</td>
<td>High: Follow up: 135/161 (83.9%). A large amount of missing data in intervention versus control. Analysis not intention-to-treat. Non-starters excluded from analysis. No data on characteristics of non-starters. Data imputed by last observation carried forward for drop outs. Reasons for non-starters and drop outs given.</td>
<td>Low: All listed outcomes reported. Trial registered on ANZCTR</td>
<td>High</td>
</tr>
<tr>
<td>Trial</td>
<td>Sequence generation</td>
<td>Allocation concealment</td>
<td>Blinding of outcome assessor/analysis</td>
<td>Incomplete outcome data</td>
<td>Selective outcome reporting</td>
<td>Overall risk of bias</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>--------------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>----------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Huang 2011 (27) (Depressed postnatal population)</td>
<td>Low: a randomised table used by the researcher</td>
<td>High: Randomised list visible to researcher</td>
<td>Unclear: Outcome assessment blinded but unclear if researchers conducting analysis blinded</td>
<td>High: Follow up (postnatal and control) 128/160 (80%). Missing data identical in both arms. Analysis not intention-to-treat. Drop outs not included in analysis.Completers and drop outs were not statistically different in demographics.</td>
<td>Unclear: All listed outcomes reported (weight reported but not BMI). Trial not registered on trial registries (proposed outcomes unavailable)</td>
<td>High</td>
</tr>
<tr>
<td>Surkan 2012 (30) (General postnatal population)</td>
<td>Unclear: Stratified randomisation by BMI (&lt;29, &gt;=29) and trial region. No further details of how sequence was generated</td>
<td>Unclear: No details</td>
<td>Unclear: Outcome assessor blinded. Unclear if researcher conducting analysis blinded</td>
<td>High: Follow up: 403/679 (59.35%). Analysis was not Intention-to-treat. Only 322/679 included in analysis. Drop outs and those not receiving ‘minimal intervention’ excluded. Proportion breastfeeding significantly lower among drop outs. Reasons for non-completion given.</td>
<td>Unclear: Main outcome reported. No results on other primary outcomes listed in methods paper. Trial not registered on trial registries (proposed outcomes unavailable)</td>
<td>High</td>
</tr>
<tr>
<td>Trial</td>
<td>Sequence generation</td>
<td>Allocation concealment</td>
<td>Blinding of outcome assessor/analysis</td>
<td>Incomplete outcome data</td>
<td>Selective outcome reporting</td>
<td>Overall risk of bias</td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>--------------------------------------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Thiruppathi 2014 (33) (General postnatal population)</td>
<td>Unclear: No details</td>
<td>Unclear: No details</td>
<td>Unclear: No details</td>
<td>High: Follow up 41/45 (91.1%) participants included in analysis. Not intention to treat analysis. Two mothers who did not receive the intervention/ control excluded from analysis. Number of drop outs and reasons for dropping out balanced across arms. No detail provided on differences between those included/ lost to follow up.</td>
<td>Unclear: Listed primary outcome measures reported on. Trial not registered on trial registries (proposed outcomes unavailable)</td>
<td>High</td>
</tr>
<tr>
<td>Trial</td>
<td>Sequence generation</td>
<td>Allocation concealment</td>
<td>Blinding of outcome assessor/analysis</td>
<td>Incomplete outcome data</td>
<td>Selective outcome reporting</td>
<td>Overall risk of bias</td>
</tr>
<tr>
<td>------------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>--------------------------------------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
| Keller 2014 (34) (General postnatal population) | Unclear: No details | Unclear: No details | Unclear: No details | High: Follow up 93/139 (66.9%)  
Loss to follow up not balanced between arms:  
Intervention follow up 54/68: 79.4%  
Comparator follow up 39/71: 54.9%  
Reasons for loss to follow up were provided.  
No imputation  
Not intention to treat analysis – drop outs excluded | Low: All listed outcomes reported. Trial registered on clinical trials.gov | High |
| Lewis 2014 (32) (General postnatal population) | Low: Random number table used, with permuted blocks of varying sizes | Unclear: No details | Low: Follow up blinded to group allocation.  
Comparator was a wellness/support intervention, participants received the same level of contact as those in the exercise intervention. | Low: 124/130 (95.4%) participants included in the analysis.  
Not intention-to-treat analysis as drop outs excluded.  
Number of drop outs not balanced between arms (5 lost in intervention, 1 in comparator). No reasons given for drop outs.  
Differences between drop outs and those followed up not presented. | Unclear: Trial registered on clinical trials.gov NCT00961402. All outcome measures recorded have been reported in the results paper. Baseline EPDS not provided. PHQ-9 medians provided but not IQRs | Unclear |
<table>
<thead>
<tr>
<th>Trial</th>
<th>Sequence generation</th>
<th>Allocation concealment</th>
<th>Blinding of outcome assessor/analysis</th>
<th>Incomplete outcome data</th>
<th>Selective outcome reporting</th>
<th>Overall risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boath 2015 (37)</td>
<td>Low: Computer generated random number table</td>
<td>Unclear: No details</td>
<td>Unclear: Randomisation was obscured from those conducting follow up, unclear if obscured from those doing analysis</td>
<td>Unclear: All those randomised included in follow up. Unclear if intention to treat analysis</td>
<td>Unclear: Listed primary outcome measures reported on. Trial not registered on trial registries (proposed outcomes unavailable)</td>
<td>Unclear</td>
</tr>
<tr>
<td>Daley 2015 (36)</td>
<td>Low: Computerised randomisation service</td>
<td>Low: Internet based randomisation: allocation concealed until randomisation took place</td>
<td>High: Outcome assessors and those conducting analysis not blinded</td>
<td>Low: Follow up overall 85/94 (90.4%) Balanced between arms: Intervention follow up 43/47 (91.5%) Comparator follow up 42/47 (89.4%) Reasons for loss to follow up reported. No Imputation Not intention to treat analysis – drop outs excluded</td>
<td>Low: All outcomes in protocol reported. Registered on ISRCTN</td>
<td>High</td>
</tr>
</tbody>
</table>

Low: Low risk of bias, High: High risk of bias, Unclear: Unclear risk of bias
BMI: Body Mass Index;
ISRCTN: International Standard Randomised Controlled Trial Number (Registry), ANZCTR: Australian New Zealand Clinical Trials Registry
EPDS: Edinburgh Postnatal Depression Questionnaire, PHQ-9: The Patient Health Questionnaire-9, PPQ: The Lederman Postpartum Self-Evaluation Questionnaire
APPENDIX 3: Funnel plot of the distribution of publications
APPENDIX 4: Meta-analysis of the effect of exercise on depressive symptoms in depressed and general postnatal populations

(Standardised mean difference)

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Exercise</th>
<th>Comparator</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
<td>Mean</td>
</tr>
<tr>
<td>Armstrong 2003</td>
<td>-1.29</td>
<td>3.76</td>
<td>10</td>
<td>-3.7</td>
</tr>
<tr>
<td>Robichaud 2008</td>
<td>-1.68</td>
<td>3.74</td>
<td>25</td>
<td>-0.48</td>
</tr>
<tr>
<td>Daley 2008</td>
<td>-4.8</td>
<td>4.81</td>
<td>12</td>
<td>-4.5</td>
</tr>
<tr>
<td>Da Costa 2009</td>
<td>-5.39</td>
<td>4.88</td>
<td>46</td>
<td>-5.2</td>
</tr>
<tr>
<td>Hsing 2011</td>
<td>3.01</td>
<td>5.87</td>
<td>64</td>
<td>4.5</td>
</tr>
<tr>
<td>Daley 2015</td>
<td>-5.28</td>
<td>4.24</td>
<td>47</td>
<td>-4.95</td>
</tr>
<tr>
<td>Doch 2015</td>
<td>-0.9</td>
<td>5.52</td>
<td>11</td>
<td>-3.2</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>214</td>
<td>203</td>
<td>50.8%</td>
<td>-0.32 [-0.63, -0.00]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.08, Chi² = 13.41, df = 6 (P = 0.04), I² = 55%
Test for overall effect: Z = 1.87 (P = 0.05)

2.1.2 General postnatal population

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Exercise</th>
<th>Comparator</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
<td>Mean</td>
</tr>
<tr>
<td>Norman 2010</td>
<td>-3.27</td>
<td>5.17</td>
<td>62</td>
<td>-0.21</td>
</tr>
<tr>
<td>Surkan 2012</td>
<td>-1.3</td>
<td>9.06</td>
<td>203</td>
<td>1.23</td>
</tr>
<tr>
<td>Haman 2013</td>
<td>-0.5</td>
<td>3.87</td>
<td>40</td>
<td>-1.8</td>
</tr>
<tr>
<td>Thimmapathi 2014</td>
<td>-3</td>
<td>0.86</td>
<td>28</td>
<td>-0.24</td>
</tr>
<tr>
<td>Keller 2014</td>
<td>-1.2</td>
<td>4.3</td>
<td>38</td>
<td>-1.71</td>
</tr>
<tr>
<td>Lavoie 2014</td>
<td>-2.08</td>
<td>2.75</td>
<td>81</td>
<td>-1.87</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>433</td>
<td>458</td>
<td>49.2%</td>
<td>-0.57 [-1.12, -0.02]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.41, Chi² = 65.74, df = 5 (P < 0.00001), I² = 92%
Test for overall effect: Z = 2.04 (P = 0.04)

Total (95% CI)

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td>646</td>
<td>661</td>
</tr>
</tbody>
</table>

100.0%  -0.44 [-0.75, -0.12]

Heterogeneity: Tau² = 0.25, Chi² = 79.17, df = 12 (P < 0.00001), I² = 86%
Test for overall effect: Z = 2.72 (P = 0.006)

Test for subgroup differences: Chi² = 0.92, df = 1 (P = 0.43), I² = 0%
## APPENDIX 5: Meta-analysis of the effects of different intervention types on depressive symptoms

(Standardised mean difference)

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Exercise</th>
<th>Comparator</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD Total</td>
<td>Mean SD Total</td>
<td>Weight (IV Random, 95% CI)</td>
</tr>
<tr>
<td>2.2.1 Exercise only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daley 2008</td>
<td>-4.8</td>
<td>4.31 18</td>
<td>-4.5 4.8 15</td>
<td>8.7%</td>
</tr>
<tr>
<td>Robichaud 2008</td>
<td>-1.68</td>
<td>3.74 25</td>
<td>-0.48 3.11 23</td>
<td>7.6%</td>
</tr>
<tr>
<td>Da Costa 2009</td>
<td>-5.96</td>
<td>4.24 48</td>
<td>-5.2 4.04 42</td>
<td>8.6%</td>
</tr>
<tr>
<td>Hara 2013</td>
<td>-0.5</td>
<td>3.67 48</td>
<td>-1.8 3.24 47</td>
<td>8.7%</td>
</tr>
<tr>
<td>Lexin 2014</td>
<td>-2.08</td>
<td>2.75 81</td>
<td>-1.37 3.97 63</td>
<td>9.6%</td>
</tr>
<tr>
<td>Thuoapathy 2014</td>
<td>-3.04</td>
<td>0.64 20</td>
<td>-0.24 0.52 21</td>
<td>4.6%</td>
</tr>
<tr>
<td>Daley 2015</td>
<td>-5.28</td>
<td>4.24 41</td>
<td>-4.05 4.18 38</td>
<td>8.5%</td>
</tr>
<tr>
<td>Boleh 2015</td>
<td>-8.9</td>
<td>5.52 11</td>
<td>-3.2 4.68 11</td>
<td>5.5%</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>268</td>
<td>260</td>
<td>51.7%</td>
<td>-0.56 [-1.13, 0.01]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.06; Chi² = 84.03, df = 7 (P < 0.00001), I² = 69%
Test for overall effect Z = 1.91 (P = 0.05)

2.2.2 Exercise co-intervention

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Exercise</th>
<th>Comparator</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armstrong 2003</td>
<td>-12.8</td>
<td>3.76 10</td>
<td>-3.7 6.13 10</td>
<td>4.7%</td>
</tr>
<tr>
<td>Norman 2010</td>
<td>-3.27</td>
<td>5.17 62</td>
<td>-0.21 4.04 13</td>
<td>9.1%</td>
</tr>
<tr>
<td>Hing 2011</td>
<td>3.01</td>
<td>5.87 64</td>
<td>4.5 7.02 64</td>
<td>9.6%</td>
</tr>
<tr>
<td>Sapon 2012</td>
<td>-1.3</td>
<td>9.86 203</td>
<td>-1.2 10.9 203</td>
<td>9.8%</td>
</tr>
<tr>
<td>Keeler 2014</td>
<td>-1.21</td>
<td>4.39 38</td>
<td>-1.71 4.04 64</td>
<td>8.1%</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>378</td>
<td>401</td>
<td>41.3%</td>
<td>-0.35 [-0.68, 0.04]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.06; Chi² = 14.03, df = 4 (P = 0.017), I² = 72%
Test for overall effect Z = 2.19 (P = 0.03)

Total (95% CI)

646     661 100.0% 0.04 [0.75, 0.12]

Heterogeneity: Tau² = 0.26; Chi² = 78.17, df = 12 (P < 0.00001); I² = 56%
Test for overall effect Z = 2.72 (P = 0.000)
Test for subdomain differences: Chi² = 0.41, df = 1 (P = 0.52), I² = 0.0%
### APPENDIX 6: Meta-analysis of the effects of different exercise contexts on depressive symptoms

(Standardised mean difference)

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Exercise</th>
<th>Comparator</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD Total</td>
<td>Mean</td>
<td>SD Total</td>
</tr>
<tr>
<td>2.3.1 Group exercise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armstrong 2003</td>
<td>-12.8</td>
<td>3.78 10</td>
<td>-3.7</td>
<td>6.13 10</td>
</tr>
<tr>
<td>Nunnally 2010</td>
<td>-3.27</td>
<td>5.17 62</td>
<td>-0.21</td>
<td>4.94 73</td>
</tr>
<tr>
<td>Hanraha 2013</td>
<td>-0.5</td>
<td>3.07 48</td>
<td>-1.8</td>
<td>3.24 47</td>
</tr>
<tr>
<td>Inkin 2014</td>
<td>-1.21</td>
<td>4.3 38</td>
<td>-1.71</td>
<td>4.04 54</td>
</tr>
<tr>
<td>Thiappanpilli 2014</td>
<td>3.64</td>
<td>20 0.24 21</td>
<td>4.0</td>
<td>4.68 11</td>
</tr>
<tr>
<td>Bocchi 2015</td>
<td>-8.9</td>
<td>5.52 11</td>
<td>-3.2</td>
<td>4.88 11</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>180</td>
<td>246 48.7%</td>
<td>-1.10</td>
<td>[1.95, 4.21]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 1.08; Chi² = 75.38, df = 6 (P < 0.00001); I² = 93%

Test for overall effect Z = 2.62 (P = 0.01)

2.3.2 Participant choice

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Comparator</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD Total</td>
<td>Mean</td>
</tr>
<tr>
<td>Dailly 2008</td>
<td>-4.8</td>
<td>4.91 18</td>
<td>-4.5</td>
</tr>
<tr>
<td>Fintochoud 2008</td>
<td>-1.38</td>
<td>2.74 25</td>
<td>-1.49</td>
</tr>
<tr>
<td>De Costa 2009</td>
<td>-5.5</td>
<td>3.98 40</td>
<td>-5.2</td>
</tr>
<tr>
<td>Huping 2011</td>
<td>3.01</td>
<td>5.97 64</td>
<td>4.5</td>
</tr>
<tr>
<td>Srikan 2012</td>
<td>1.06</td>
<td>9.06 203</td>
<td>1.23</td>
</tr>
<tr>
<td>Laxie 2014</td>
<td>-2.08</td>
<td>2.75 61</td>
<td>-1.37</td>
</tr>
<tr>
<td>Dailly 2015</td>
<td>5.29</td>
<td>4.24 41</td>
<td>4.95</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>1456</td>
<td>446 55.3%</td>
<td>0.20 [-0.33, -0.08]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.00; Chi² = 2.32, df = 6 (P = 0.90); I² = 0%

Test for overall effect Z = 2.32 (P = 0.01)

Total (95% CI) | 046 | 661 100.0% | -0.41 [-0.75, 0.12] |

Heterogeneity: Tau² = 0.25; Chi² = 79.17, df = 12 (P < 0.00001); I² = 95%

Test for overall effect Z = 2.72 (P = 0.0066)

Test for subnational differences: Chi² = 0.49, df = 1 (P = 0.05), I² = 74.9%