The effectiveness of exercise for the prevention and treatment of antenatal depression: systematic review with meta-analysis

AJ Daley,a L Foster,b,* G Long,b,* C Palmer,b,* O Robinson,b,* H Walmsley,b,* R Wardb,*

a Primary Care Clinical Sciences, b College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

Correspondence: Dr AJ Daley, Primary Care Clinical Sciences, University of Birmingham, Birmingham, UK B15 2TT.
Email a.daley@bham.ac.uk

Accepted 5 May 2014. Published Online 17 June 2014.

Background Antenatal depression can have harmful consequences for the mother and fetus. Exercise may be a useful intervention to prevent and treat antenatal depression.

Objectives This systematic review aims to establish whether there is sufficient evidence to conclude that exercise is an effective intervention for preventing and treating antenatal depression.

Search strategy Searches using electronic databases from MEDLINE, Cochrane Library, CINAHL, EMBASE, AMED and PsycINFO were performed.

Selection criteria Randomised controlled trials (RCT) that compared any type of exercise intervention with any comparator in pregnant women were eligible for inclusion.

Data collection and analysis Meta-analysis was performed calculating standardised mean differences (SMD).

Main results Six trials (seven comparisons) were eligible for inclusion. Meta-analysis showed a significant reduction in depression scores (SMD –0.46, 95% CI –0.87 to –0.05, P = 0.03, I² = 68%) for exercise interventions relative to comparator groups. The test for subgroup differences in women who were non-depressed (one trial) (SMD –0.74, 95%CI –1.22 to –0.27, P = 0.002) and depressed (five trials) (SMD –0.41, 95% CI –0.88 to 0.07, P = 0.09) at baseline was not significant (P = 0.32). The test for subgroup differences between aerobic (one trial) and non-aerobic exercise (five trials) was also nonsignificant (P = 0.32).

Authors’ conclusions We found some evidence that exercise may be effective in treating depression during pregnancy but this conclusion is based on a small number of low-moderate quality trials with significant heterogeneity and wide confidence intervals.

Keywords Antenatal, depression, exercise, meta-analysis, systematic review.

Linked article This article is commented on by G Shivakumar, p. 63 in this issue. To view this mini commentary visit http://dx.doi.org/10.1111/1471-0528.13053.

Introduction Antenatal depression is a common mental illness with a high prevalence rate affecting between 10 and 13% of pregnancies.1 It is associated with harmful consequences to both the mother and child, including premature labour,2,3 low birthweight and a compromised mother–child relationship.4,5 Women who have previously been depressed are most at risk for experiencing antenatal depression.6 The substantial risks to the mother and baby from antenatal depression emphasises the importance of finding effective interventions. Current treatment options include antidepressant medication and psychological therapies.7,8 Antidepressants are rarely prescribed due to concerns of adverse effects on the fetus and its neonatal development.9,10 Psychotherapy, such as cognitive behavioural therapy (CBT), has also been shown to be effective in treating depression in pregnant women11 but there can be long waiting lists to access these interventions, and for many women it is not affordable.12,13

Another treatment option that has been proposed is exercise. In the absence of medical or obstetric complications, the Royal College of Obstetricians and Gynaecologists in the UK, and American College of Obstetricians and

*These authors conducted this study as part of their undergraduate medical degree training and contributed equally to generating the manuscript and therefore are listed alphabetically.
Gynecologists recommend 30 min of aerobic exercise per day to obtain both physical and mental health benefits during pregnancy. In 2007, NICE published a guideline that exercise should be an option for the management of antenatal depression or mild depression that develops during pregnancy. However, this guideline was based on studies from general populations and there was a lack of evidence pertaining to pregnant women per se. A recent narrative review of exercise for antenatal depression summarised data from six small observational studies and found that exercise was associated with better mental health outcomes during pregnancy, but to date no published meta-analysis has evaluated the effectiveness of exercise specifically for the prevention and treatment of antenatal depression. This study aims to review systematically the available evidence from randomised controlled trials (RCTs) on the effectiveness of exercise as an intervention in the prevention and treatment of antenatal depression and, where possible, perform meta-analyses.

**Methods**

**Search strategy for identification of studies**
The electronic databases searched (and dates) were as follows: Cochrane Library CENTRAL Register of Controlled Trials Issue 1 of 12 2014, CDSR Issue 1 of 12 2014, DARE and HTA Issue 1 of 4, 2014; MEDLINE(Ovid) In-Process & Other Non-Indexed Citations February 4, 2014; MEDLINE (Ovid) 1946 to January Week 4, 2014; Embase (Ovid) 1980 to 2014 Week 05; AMED Allied and Complementary Medicine Database (Ebsco) 1995 – 4 February 2014; PsycINFO (Ovid) 1967 to January Week 3, 2014. An appropriate combination of text words and index terms was used to describe population, condition and intervention. A sample MEDLINE search strategy used to identify trials for this review is included as Table S1; this was adapted for use on each database. Searches were conducted by an information specialist (S.B.) at the University of Birmingham between 5 and 7 February 2014 and were not restricted by date or language. The bibliographies of studies were also searched to identify additional studies. Relevant review articles were evaluated for information on additional trials. Information regarding on-going research trials was obtained by searching the WHO ICTRP (International Clinical Trials Registry Platform) and ClinicalTrials.gov.

**Study selection**
This review included RCTs that compared any type of exercise intervention with usual care, control groups or any other comparator. Trials involving exercise as a co-intervention were also eligible. This review focused on both the prevention and treatment of antenatal depression and therefore any trials that had recruited non-depressed pregnant women or women at risk or diagnosed with antenatal depression were eligible. Trials were excluded if the intervention lasted <6 weeks or did not provide data on depression antenatally.

**Data extraction**
Two reviewers (G.L. and R.W.) initially screened the title and abstracts of identified studies. Full articles of potentially relevant studies were retrieved for detailed evaluation by all authors. Three authors (A.D., L.F. and C.P.) conducted the final selection of studies to be included in the review using a standardised eligibility form. If the study fulfilled the review inclusion criteria, data was extracted by two reviewers (O.R. and H.W.) using a standard extraction form. The data including study design, participant characteristics, number of study participants and the number completing follow-up antenatally, recruitment procedures, length and content of the exercise intervention, intervention adherence, type of comparator and outcome assessments were extracted by two reviewers (O.R. and H.W.) using a standard extraction form. Reviewers were not blinded to journal of publication, author names or their institutions. Authors of studies were contacted if additional information was required adequately to complete the data extraction form. Refer to Table S2 for a summary of included studies.

**Assessment of study quality**
Two reviewers (L.F. and C.P.) independently assessed the methodological quality of each study. The scoring system was modified from the Delphi List Criteria, which is a set of nine criteria for quality assessment of RCTs (Table S3), tailored to exercise-based research. In this systematic review, trials were assessed against seven of the nine Delphi List quality criteria. The criteria relating to the use of blinding were not rated (i.e. blinding of the care providers and blinding of the patients) because it is very hard, if not impossible, to conduct exercise intervention trials where patients and care providers are blinded to the intervention. Blinding of the outcome assessor, however, was included. Discrepancies concerning the quality of a particular trial were resolved by consensus with a third reviewer (A.D.).

**Data synthesis and data analysis**
The primary outcome was change in depression score between baseline and final antenatal follow-up using any form of assessment. Means and standard deviation (SD) of the difference scores for depression were extracted from study reports by one reviewer (A.D.), or by conversion of standard errors and other measures of variability. Where the SD of the difference score was not reported, this was calculated. We used a correlation of 0.6 in our assumptions to estimate SD differences which was based on data from a previous trial of exercise during pregnancy. Standardised mean difference (SMD) was calculated to summarise effects...
from trials in the meta-analysis. For as much of the available data to be included in the primary analyses as possible, all trials were pooled irrespective of the population recruited or the type of exercise intervention tested. The final assessment of outcomes was used when studies reported multiple follow-up times.

A fixed effect meta-analysis was undertaken in the absence of heterogeneity, otherwise a random effects model was used. The $I^2$ statistic was used to assess heterogeneity, with $I^2 > 50\%$ considered important. Subgroup analyses were also performed for populations of women (non-depressed versus depressed/at risk of depression at baseline) and type of exercise intervention (aerobic versus non-aerobic) (i.e. yoga, tai chi). One of the included trials included three experimental groups (yoga, massage and usual care); data from these groups were entered in the meta-analysis as two separate comparisons (i.e. yoga versus usual care and yoga versus massage). To avoid counting the same women in the meta-analysis twice, the sample size in the yoga group was halved ($n = 11$) in both comparisons. Statistical analyses were performed using REVAMAN 5.2 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark).

Results

Trial flow

A total of 919 studies were identified from the search strategy and six trials (seven comparisons) were eligible for inclusion (see Table S2). Of these six trials, one was found through the bibliography of one of the sources. Trial flow is shown in Figure 1. The main reasons for exclusion of full text articles assessed for eligibility were no measure of depression, not an RCT, and compared two exercise interventions.

Characteristics of included studies

Table S2 shows the characteristics of the six trials, which recruited 406 pregnant women of which 348 (85.7%) completed follow up. Studies were published between 2008 and 2013 in the USA and Colombia. Four studies had similar investigators. Women were recruited from 16 weeks’ gestation and ranged from 14 to 38 years old. One study included non-depressed women and five included women at risk of depression/depressed at baseline. In the study of women at risk of depression, 35% had reported being treated for a psychiatric disorder in the past, 32% had taken psychotropic medication in the past, 52% exceeded a score of 14 on the perceived stress scale at baseline, and 31% exceeded a score of 16 on the Center for Epidemiologic Studies Depression Scale (CES-D). This trial is therefore considered to have recruited depressed women. One trial evaluated an aerobic-based exercise intervention and five trials evaluated non-aerobic-based exercise.

Activities that depend primarily on the aerobic energy generating process, such as walking, running, swimming and aerobics classes, were considered to be aerobic exercise. Strength training and activities such as yoga and tai chi were defined as non aerobic exercise. Interventions lasted eight and 12 weeks. Standard prenatal care was the comparator in two of trials, two studies used waiting-list controls, one used parenting education sessions and one used social support. One trial had two comparison groups (massage and a control group). All trials evaluated exercise as single intervention and used the Centre for Epidemiological Studies-depression scale (CES-D) to measure depression. One study used both the CES-D and the Edinburgh Postnatal Depression Scale antenatally; data from the CES-D was used in the meta-analysis. Only one trial provided follow-up data postnatally.

Methodological quality of included studies

Study quality results ranged from three to six points, of a maximum of seven points (Table S3). All studies specified eligibility criteria, conducted randomisation and produced point estimates and measures of variability for the primary outcome. Groups were similar at baseline in six trials and one trial did not report any information to make this assessment. Two studies clearly indicated that blinding of outcome assessors had occurred and one indicated concealed treatment allocation had occurred. None of the studies stated that intention-to-treat analysis was used.
Meta-analysis results

Six RCTs (seven comparisons) were eligible for inclusion in the meta-analysis, which showed a significant reduction in depression scores (SMD = −0.46, 95% CI −0.87 to −0.05, \( P = 0.03, I^2 = 68\% \)) for exercise relative to comparators (see Figure 2). The test for subgroup differences in women who were non-depressed (one RCT) \( (SMD = −0.74; 95\% CI −1.22 \text{ to } −0.27, P = 0.002) \) and depressed (five trials) \( (SMD = −0.41; 95\% CI −0.88 \text{ to } 0.07, P = 0.09, I^2 = 70\% \) at baseline was nonsignificant \( (P = 0.32, I^2 = 68\%) \). The test for subgroup differences between aerobic (one trial) \( (SMD = −0.74; 95\% CI −1.22 \text{ to } −0.27, P = 0.002) \) and non-aerobic exercise (five trials) \( (SMD = −0.41; 95\% CI −0.88 \text{ to } 0.07, P = 0.09, I^2 = 70\%) \) was nonsignificant \( (P = 0.32, I^2 = 68\%) \). Significant heterogeneity was present in the primary analysis of exercise versus comparators and subgroup analyses of depressed women and non-aerobic exercise (not estimated for the non-depressed and aerobic exercise, as only one trial included in these subgroups). No trials reported data on safety or adverse events. Funnel plots were not used to investigate the possibility of publication bias because of the small number of trials included.

Adherence to exercise interventions

Attendance at exercise classes or adherence to physical intervention guidelines was reported in two of the trials. One trial \( (24) \) reported that participants received on average 28.9/36 sessions over 3 months. One trial \( (25) \) reported that women in the intervention group attended a mean number of 7.2/8 sessions and reported in weekly diaries to engage in an average of about 77 min per week in formal mindfulness practice (which included hatha yoga/walking and moving mindfulness practice). Four studies \( (20–23) \) did not provide any data on adherence or attendance rates to interventions.

Discussion

This systematic review with meta-analysis found preliminary evidence (small–moderate effect size) to suggest that exercise may be effective in reducing depression during pregnancy. Five of the six included studies recruited women with depression so it is not possible to indicate whether exercise is likely to prevent antenatal or postnatal depression although one included study did show that an aerobic exercise-based intervention was effective in preventing antenatal depression. Regardless of this, the findings of this review are based on a small number of low to moderate quality trials with significant heterogeneity and wide confidence intervals, meaning that estimates may be imprecise. The tests for subgroup difference for non-depressed versus depressed women and aerobic versus non-aerobic exercise were nonsignificant, although the subgroups for non-depressed women and aerobic exercise interventions were based on the same single trial. It would be important for future research to include larger samples to explore these subgroup questions in more depth and there is a particular need for studies to assess the merits of aerobic-based exercise interventions such as walking and aqua-natal classes.

Strengths and weakness of the review

This is the first meta-analysis to investigate the effectiveness of exercise to prevent and treat antenatal depression and therefore it has the opportunity to substantially add to knowledge. Loss to follow up was low (14.3%) overall and therefore data for the vast majority of randomised participants were included in the meta-analyses. Loss to follow up was also balanced across the groups and inspection of the Forest plot shows that the number of participants lost to follow up was not related to size of effect. We have conducted sensitivity/subgroup analyses to try and better

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Exercise Mean</th>
<th>SD</th>
<th>Total</th>
<th>Comparator Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Std. mean difference IV, Random, 95% CI</th>
<th>Std. mean difference IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field 2012 (c1)</td>
<td>−8.2</td>
<td>8.8</td>
<td>11</td>
<td>−3.4</td>
<td>8.1</td>
<td>21</td>
<td>12.7%</td>
<td>−0.56 [−1.31, 0.18]</td>
<td></td>
</tr>
<tr>
<td>Field 2012 (c2)</td>
<td>−8.2</td>
<td>8.8</td>
<td>11</td>
<td>−11.8</td>
<td>5.1</td>
<td>24</td>
<td>12.9%</td>
<td>0.55 [−0.18, 1.27]</td>
<td></td>
</tr>
<tr>
<td>Field 2013a</td>
<td>−8.9</td>
<td>8.7</td>
<td>37</td>
<td>−2.8</td>
<td>10.1</td>
<td>38</td>
<td>17.1%</td>
<td>−0.64 [−1.10, −0.18]</td>
<td></td>
</tr>
<tr>
<td>Field 2013b</td>
<td>−9.2</td>
<td>8.8</td>
<td>40</td>
<td>−9.9</td>
<td>9.9</td>
<td>39</td>
<td>17.5%</td>
<td>0.08 [−0.36, 0.52]</td>
<td></td>
</tr>
<tr>
<td>Mitchell 2012</td>
<td>−9.6</td>
<td>8.5</td>
<td>11</td>
<td>−1</td>
<td>8.4</td>
<td>11</td>
<td>10.6%</td>
<td>−0.98 [−1.87, −0.08]</td>
<td></td>
</tr>
<tr>
<td>Robledo-Colonlia 2012</td>
<td>−6.6</td>
<td>6.5</td>
<td>37</td>
<td>−1</td>
<td>6.8</td>
<td>37</td>
<td>17.0%</td>
<td>−0.74 [−1.22, −0.27]</td>
<td></td>
</tr>
<tr>
<td>Vieten 2008</td>
<td>−4.2</td>
<td>7.1</td>
<td>13</td>
<td>3</td>
<td>6</td>
<td>18</td>
<td>12.3%</td>
<td>−1.08 [−1.85, −0.31]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>160</td>
<td>188</td>
<td>100.0%</td>
<td>−0.46 [−0.87, −0.05]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: \( \text{Tau}^2 = 0.20, \text{Chi}^2 = 18.82, df = 6 (P = 0.004); P = 68\% \)
Test for overall effect: \( Z = 2.22 (P = 0.03) \)

Figure 2. Primary meta-analysis of all trials for antenatal depression scores. The comparison Field (2012) (c1) refers to yoga versus usual care. The comparison Field (2012) (c2) refers to yoga versus massage.
understand the relationship between exercise and antenatal depression. This study provides evidence that may be used to help strengthen and direct clinical guidance and assist healthcare professionals in their decision-making regarding exercise as a potential treatment for antenatal depression.

The findings of this review are based on a small sample size therefore caution is required when interpreting the results. Although participants in all the included studies completed follow up at the end of the intervention period, participants were at varying weeks of gestation when follow up took place, ranging from 28 to 41 weeks’ gestation. Various types of exercise were tested, ranging from moderate intensity aerobic exercise to forms of exercise requiring low levels of physical exertion, making it difficult to comment on what type of exercise might be most effective. It may be the case that an intervention dose effect exists, but we were not able to explore this question due to a lack of studies. This would be an important question for future research to explore. It is also important to note that all of the trials in the analysis of women who were depressed at baseline evaluated predominately yoga-based activities.

Interpretation
Guidance from NICE\textsuperscript{16} states that health professionals should consider exercise a treatment for antenatal depression, and the Royal College of Obstetricians and Gynaecologists\textsuperscript{14} and the American College of Obstetricians and Gynecologists\textsuperscript{15} have stated that exercise can provide mental health benefits during pregnancy, but more evidence is required to qualify this guidance. This review now provides some evidence to support clinical guidance and recommendations about exercise as a treatment for antenatal depression. Our findings are also consistent with a meta-analysis of the effectiveness of exercise as a treatment for postnatal depression.\textsuperscript{28} It remains unknown whether exercise prevents antenatal depression. We are aware of one prevention-based trial\textsuperscript{29} that recruited women (\textit{n} = 855) antenatally and evaluated a 12-week exercise intervention (aerobic and strengthening exercise) and assessed depression 3 months postnatally. This trial did not report a significant difference between the exercise and usual care groups in the prevalence of depression postnatally, although this study may have recruited an atypical sample as prevalence was very low (2–4%) in both groups at follow up. This trial was excluded from this review because it did not report a measure of depression antenatally.

Conclusions
We found some evidence to suggest that exercise may be effective in treating depression during pregnancy but this conclusion is based on a small number of low–moderate quality trials with significant heterogeneity and wide confidence intervals. No conclusions can be made about exercise in preventing antenatal depression due to a lack of trials. More high quality trials that address these questions are now required.

Note
In completing this review we noted some reporting errors and missing information in three of the included studies [Field et al.\textsuperscript{20,22}; Mitchell et al.\textsuperscript{23}], which we have clarified with the authors. Therefore our reporting of these trials may differ from the information in these published reports.

In Field et al.\textsuperscript{20} the number of participants providing follow up in each experimental condition was not reported. The authors provided us with this information to allow us to calculate loss to follow up per group.

In Field et al.\textsuperscript{22} incorrect data for participants in the yoga group was reported in their Table 1 and the study authors provided us with the correct data. The trial flow diagram did not specify which boxes referred to the yoga and social support groups respectively and the authors provided us with this information to allow us to calculate loss to follow up per group.

In Mitchell et al.\textsuperscript{23} no data regarding how many randomised participants (\textit{n} = 24) completed follow up for each group are reported. We contacted the second author of the report who confirmed 11 participants per group completed follow up.

Disclosure of interests
None.

Contribution to authorship
AD developed the initial idea for the study. SB conducted the searches. GL and RW initially screened the title and abstracts of identified studies. LF and CP independently assessed the methodological quality of each included study. AD, LF and CP conducted the final selection of studies to be included in the review. OR and HW extracted information and data from the studies. AD conducted the analyses. All authors have read and approved the final manuscript.

Details of ethics approval
Not required.

Funding
None.

Acknowledgements
We would like to thank Sue Bayliss (S.B.) for conducting the searches for this review. We would like to thank Tiffany Field for responding to our queries about her studies.
Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Search strategy.
Table S2. Characteristics of included studies.
Table S3. Table of study quality.

References


Daley et al.