Archives of Women's Mental Health (2019) 22:37–53
https://doi.org/10.1007/s00737-018-0869-3

REVIEW ARTICLE

The effectiveness of exercise-based interventions for preventing or treating postpartum depression: a systematic review and meta-analysis

Tim Carter[1] · Anastasios Bastounis[1] · Boliang Guo[1] · C Jane Morrell[1]

Received: 31 January 2018 / Accepted: 24 May 2018 / Published online: 7 June 2018
© Springer-Verlag GmbH Austria, part of Springer Nature 2018

Abstract

Postpartum depression can have detrimental effects on both a mother's physical and mental health and on her child's growth and emotional development. The aim of this study is to assess the effectiveness of exercise/physical activity-based interventions in preventing and treating postpartum depressive symptoms in primiparous and multiparous women to the end of the postnatal period at 52 weeks postpartum. Electronic databases were searched for published and unpublished randomised controlled trials of exercise/physical activity-based interventions in preventing and treating depressive symptoms and increasing health-related quality of life in women from 4 to 52 weeks postpartum. The results of the studies were meta-analysed and effect sizes with confidence intervals were calculated. The Grading of Recommendations Assessment and Development and Evaluation (GRADE) system was used to determine the confidence in the effect estimates. Eighteen trials conducted across a range of countries met the inclusion criteria. Most of the exercise interventions were aerobic and coaching compared to usual care, non-intervention and active controls. Small effect sizes of exercise-based interventions in reducing depressive symptoms were observed collectively and the quality of evidence was low across the individual studies. Although exercise-based interventions could create an alternative therapeutic approach for preventing major depression in postpartum women who experience sub-threshold elevated depressive symptoms, the clinical effectiveness and the cost-effectiveness of exercise-based and physical activity interventions need to be better established. There is a need for further more rigorous testing of such interventions in high-quality randomised controlled trials against active control conditions before large-scale roll-out of these interventions in clinical practice is proposed.

Keywords Exercise · Physical activity · Postpartum depression · Review · Meta-analysis

Introduction

About 20% of women globally experience a perinatal mental health disorder, mainly depression and anxiety, when they are pregnant or in the perinatal period up to 52 weeks after they have given birth (WHO 2017). The most severely affected women can develop self-harm and suicidal ideations (Pope et al. 2013; Wisner et al. 2013). Perinatal anxiety and depression can compromise the long-term growth and development of the baby (Farías-Antúnez et al. 2017), with long-term costs of £8.1bn (Bauer et al. 2014). A range of physical, genetic and socioeconomic factors put pregnant and postpartum women at risk of perinatal mental health problems whilst buffering factors (e.g. supportive partner) are protective (Austin et al. 2010).

In the UK, early psychosocial or pharmacological interventions are recommended to reduce the prevalence of perinatal anxiety and depression, benefit women and families and reduce costs (Morrell et al., 2009a, b; NICE 2014; Morrell et al. 2016; Saligheh et al. 2017).

The perinatal period is also characterised by difficulty in managing weight and engaging in physical activity (Gaston and Cramp, 2011). A reduction in physical activity/exercise throughout pregnancy can lead to lower self-ratings of quality of life (Campolong 2017) and can have detrimental effects on

✉ Tim Carter
timothy.carter@nottingham.ac.uk

[1] University of Nottingham, Nottingham, UK

physical health (Fazzi, Saunders, Linton, Norman, & Reynolds 2017). Sedentary behaviours have been associated with increased risk for postnatal depressive symptoms, whereas physical activity in pregnancy and postnatally has been associated with decreased risk for developing depressive symptoms (Claesson, Klein, Sydsjo, & Josefsson 2014; Teychenne & York 2013). Given that engaging in sedentary behaviours during pregnancy can be continued postpartum, exercise-based interventions could yield multi-tiered benefits for the physical and mental health of perinatal women.

Small to moderate effects on depression symptoms have been found from exercise-based interventions in adults and young people (standardised mean difference [SMD] − 0.62, 95% confidence interval [CI] − 0.81, − 0.42), compared to control conditions (Cooney et al. 2013; Carter et al. 2016). In postpartum populations, there is a promising evidence base for exercise-based interventions in preventing and treating depressive symptoms (McCurdy et al. 2017; Poyatos-León et al. 2017). The content of these interventions covers aerobic activities, stretching, yoga and exercise-based coaching. In randomised controlled trials (RCTs), exercise-based interventions have been compared to control conditions of usual care (UC) or non-intervention (NI), but few have been compared against active control (AC) or waitlist control (WLC) (Armstrong & Edwards 2004; LeCheminant et al. 2014). Most exercise-based interventions have been tested in targeted populations, such as women with elevated depression symptoms (Buttner et al. 2015) or women with a previous history of depression (Lewis et al. 2014).

There is now a need for a robust evidence synthesis that follows methodologically rigorous processes (Saligheh et al. 2017) to systematically identify the components and characteristics of interventions, and analyse their effectiveness, to promote the development of beneficial exercise-based interventions in clinical practice (Saligheh et al. 2017).

This review aims to synthesise evidence from RCTs for the clinical effectiveness of exercise-based interventions compared to all types of control in preventing and treating depressive symptoms in primiparous and multiparous women from the possible onset at 4–6 weeks postnatally (Putman et al. 2017) to the end of the postpartum period (12 months after the birth of the baby). Additionally, this review aims to identify factors associated with the effectiveness of exercise-based interventions, testing the moderating effects of the interventions: scope (universal vs. targeted); content (strongly exercise-oriented vs. exercise consulting and coaching); duration (short vs. long duration); and control condition: AC vs. UC, NI, and WLC.

Methods

The protocol of this systematic review and meta-analysis was registered with PROSPERO (Carter 2017:CRD42017068376) and the presentation of the findings conforms to PRISMA (Moher et al. 2009). The primary outcome was depression symptoms in postpartum women at post-intervention and the secondary outcomes were symptoms of anxiety and health-related quality of life (HRQoL).

Inclusion criteria

Population: primiparous or multiparous postnatal women
Intervention: exercise-based (supervised, unsupervised, coaching-based, motivational, behavioural-oriented, universal, targeted or treatment based, in a community or clinical context)
Comparison: any type of control condition (e.g. flexibility/stretching or social support sessions, UC, NI, AC, WLC)
Outcomes: depression symptoms using a validated assessment tool (e.g. Edinburgh Postnatal Depression Scale (EPDS), Patient Health Questionnaire)
Study type: published or unpublished individual RCTs or cluster RCTs

Exclusion criteria

Population: pregnant women; women with psychiatric diagnoses other than depression
Intervention: no details of the exercise component; intervention delivered before 4 weeks or after 52 weeks
Comparison: no comparison interventions were excluded
Outcomes: no depression symptom measure; outcomes before 4 weeks postpartum
Study type: non-RCTs

Search strategy

Libraries and databases searched for papers published between 1974 and June 2017 were Allied and Complementary Medicine Database (AMED), Applied Social Sciences Index and Abstracts (ASSIA), Cumulative Index to Nursing and Allied Health Literature (CINAHL), Current Controlled Trials, EMBASE (Excerpta Medica), ISRCTN Register, MEDLINE (including PubMed), National Institute for Health Research Health Technology Assessment (NIHR HTA) programme databases, PROSPERO, PsycINFO, Scopus, Science Citation Index and Conference Proceedings (Web of Science), The Cochrane Library (Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Cochrane Central Register of Controlled Trials) and World Health Organisation's International Clinical Trials Registry Platform (ICTRP). Online databases of grey literature searched were clinicaltrials.gov, International Standard Randomised

Controlled Trials Number (ISRCTN) Register, OpenGrey and ProQuest Dissertations & Theses (PQDT).

The search strategy incorporated Medical Subject Heading (MeSH) terms in five areas:

Population: Postpartum Period; and Pregnant women/ OR Postnatal care/ OR Perinatal care. Depression/ OR Depression, Postpartum/; Anxiety/ OR Anxiety Disorders/
Intervention: Exercise Test/ OR Exercise/ OR Exercise Therapy/ OR Exercise Movement Techniques/
Outcome: Depression/ OR Depression, Postpartum/; Anxiety/ OR Anxiety Disorders/
Study type: The search was optimised using the 'RCTs (plus cluster)' clinical search filter recommended by the Centre for Reviews and Dissemination (CRD 2009)

Hand-searches of public online databases and contacts with field experts were also conducted. Three syntax sets were used in combination with the MeSH terms above for searching Medline, EMBASE and PsycINFO (see Table 1).

Relevant authors were contacted when full text articles were not available, there was insufficient information provided for the inclusion criteria to be applied and there were insufficient details reported on the outcomes. Lack of reply from authors led to one study being included only in the qualitative synthesis: LeCheminant et al.. (2014).

Following initial screening of titles and abstracts, full texts of all potentially relevant studies were assessed for inclusion independently by two reviewers (TC and AB). Disagreements were resolved by discussion, or a third reviewer (JM) was consulted. Reference lists of included articles were searched for potentially eligible studies.

Data extraction

Adapted versions of the Effective Practice and Organisation of Care (EPOC) Review Group data abstraction form and the Cochrane Collaboration Form for extracting data from RCTs were used to extract data from included studies. Two reviewers (TC and AB) extracted data independently and disagreements were resolved by discussion between the two reviewers who presented their arguments to each other until agreement was made. A third reviewer (JM) would have been the final arbiter, but this process was not required at any point in this review. Extracted data included information on study authors, participant demographic characteristics, intervention and control conditions, study method, recruitment and completion rates, outcomes and measurement times, information for assessment of risk of bias and quality. Experimental conditions were coded as either (a) intervention, exercise or physical activity, yoga, coaching sessions with exercise and social support with exercise, or (b) control, UC, AC (social support sessions), NI and WLC.

Quality assessment

The quality of included studies was assessed using the Cochrane Collaboration tool for assessing risk of bias (Higgins et al. 2011). Within each specified domain, adequate reporting resulted in a rating of low risk of bias, whereas evidence of bias resulted in a rating of high risk of bias. When insufficient detail was reported for clear assessment, a rating of unclear risk of bias was given. There was also an assessment of any additional threats of bias. Two researchers (TC and AB) independently rated the risk of bias for each included study. Any disagreements were resolved after discussion. The Grading of Recommendations Assessment and

Table 1 Search syntax for Medline, EMBASE and PsychINFO

PICO heading	Syntax set
Population	postnatal.mp. OR post-natal.mp. OR perinatal.mp. OR peri-natal.mp. OR postpartum.mp. OR post-partum.mp. OR puerperium.mp. OR puerperal.mp. OR pregnan$2.mp. OR post pregnancy.mp. OR post-pregnancy.mp OR postpregnancy.mp. OR motherhood.mp. OR wom#n.mp.
Intervention	aerobic.mp. OR walking.ab. OR pram-walking.mp. OR exercise*.mp. OR (physical adj3 activity).mp. OR (physical adj3 exercise).mp. OR (exercise adj3 intervention).mp. OR exercise program$3.mp. OR yoga.mp. OR tai-chi.mp. OR taichi.mp. OR tai chi.mp. OR tai ji.mp. OR tai-ji.mp. OR (social adj3 support).mp. OR obesity.mp. OR diet.mp. OR nutrition.mp. OR mindfulness.ab. OR weight loss.mp. OR physiotherapy.ab. OR physio therapy.ab. OR physio-therapy.ab. OR fitness.mp. OR sport*.mp. OR muscle*.mp. OR stretching.mp. OR leisure.mp. OR dance.mp. OR running.mp.
Outcome	Depression.mp. OR Depressive.mp. OR Depressi$2 adj3 symptom*.mp. OR (risk* adj5 depress$3).ab. OR Anxiety adj3 symptom*.mp. OR Anxiety.mp. OR Therapy adj5 depression.mp. OR depression adj3 treatment.mp. OR Diagnosis adj3 depression.mp. OR Prevention adj3 depression.mp. OR Stress.ab. OR Mood.ab. OR Mental health.mp. OR Well-being.mp. OR Well being.mp. OR Wellbeing.mp.

Development and Evaluation (GRADE) system was used to assess confidence in the quality of evidence of individual outcomes and the strength of recommendations (Guyatt et al. 2008).

Data analysis

Data analysis was performed using RevMan Version 5.3 (Nordic Cochrane Centre 2014) and STATA Version 14 (StataCorp 2015). Standardised mean differences were computed for all included studies. Post-intervention effect sizes were computed, comparing the intervention arms of the studies to all types of control. Mean differences in the primary outcome (depression symptoms) were computed to Hedge's *g*. Hedge's *g* was obtained by subtracting control mean by intervention mean, divided by their pooled standard deviation and implementing the correction factor *J* (Borenstein et al. 2009). Given the heterogeneity of methodologically diverse studies, a random effects model was adopted. Four subgroup analyses were pre-planned and conducted: (1) universal vs. targeted interventions; (2) active exercise-orientated interventions vs. non-active exercise-orientated; (3) studies using active control groups vs. studies using other control groups; and (4) interventions of longer duration vs. interventions of shorter duration.

Results

The search yielded 20,671 abstracts following the removal of duplicates. Screening of title and abstracts resulted in 103 full text articles undergoing eligibility assessment, of which 18 were included in the review and 17 in the meta-analysis. Figure 1 presents a PRISMA Flow Chart illustrating study selection.

Table 2 presents a summary of the 18 studies included in the qualitative synthesis. Seventeen studies were included in the meta-analysis: three each from Australia (Armstrong and Edwards 2003; Armstrong and Edwards 2004; Norman et al. 2010) and the UK (Daley et al. 2008; Daley et al. 2015; Forsyth et al. 2017), six from the USA (Buttner et al. 2015; Keller et al. 2014; Lewis et al. 2014; Robichaud et al.. 2009; Shelton et al. 2015; Surkan et al. 2012) and one each from Canada (DaCosta et al. 2009), Japan (Haruna et al. 2013), Iran (Saeedi 2013), Taiwan (Yang and Chen, 2018) and India (Thiruppathi et al. 2014).

Design and sample

A RCT design was used in all 17 studies in the meta-analysis (1428 participants); five of these were pilot studies (Armstrong and Edwards 2003; Daley et al. 2008; Forsyth et al. 2017; Shelton 2015; Yang and Chen, 2018). The number of

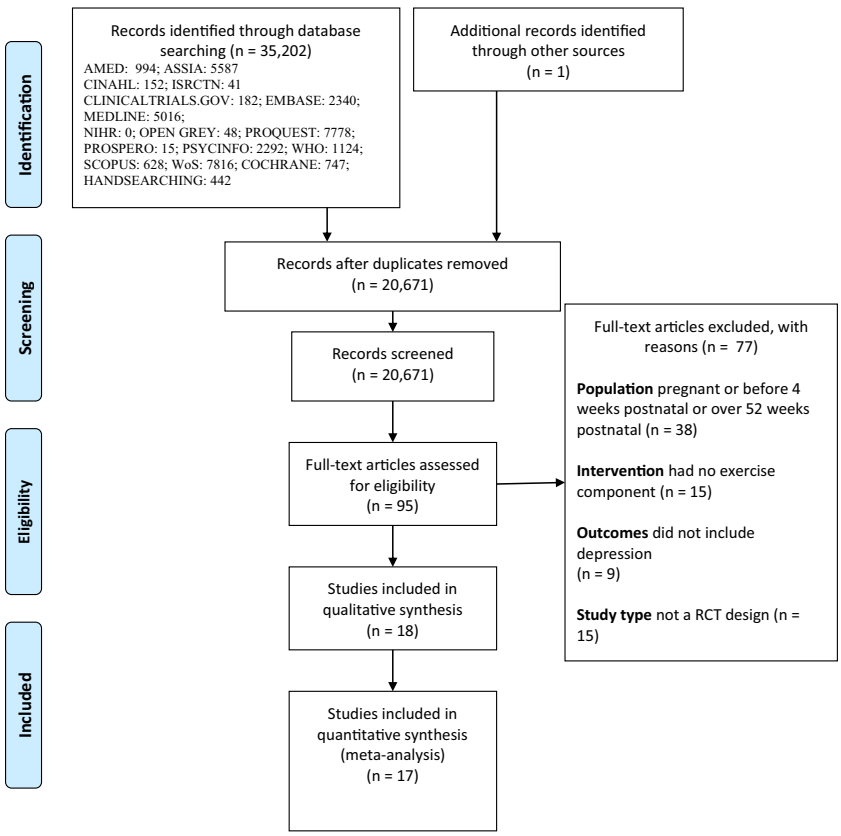

Fig. 1 PRISMA flow diagram

Table 2 Summary of study characteristics

Studies	Country	*N* (Ne; Nc)	Age range in years, (mean)	Depression inclusion criteria	Baseline depression severity[a]; baseline depressive symptoms (mean)	Measures	Assessment time points
Armstrong and Edwards (2003)	Australia	*N* = 20 (Ne = 10; Nc = 10)	Majority 21–30 (NR)	Elevated depressive symptoms (EPDS ≥ 12)	Moderate (half of the participants were taking medication for PND); *I* = 17.4; *C* = 18.4	DASS EPDS GHQ12	Baseline, week 6, week 12 (post-intervention)
Armstrong and Edwards (2004)	Australia	*N* = 19 (Ne = 9; Nc = 10)	NR	Elevated depressive symptoms (EPDS ≥ 12)	Moderate (all the participants scored ≥ 12 in EPDS and half of them were taking medication for PND); *I* = 17.25; *C* = 17.17	EPDS	Baseline, week 6, week 12 (post-intervention)
Buttner et al. (2015)	USA	*N* = 57 (Ne = 28; Nc = 29)	NR (Me = 29.81; Mc = 32.45)	Elevated depressive symptoms (HDRS ≥ 12; PHQ-9 ≥ 10)	Mild to moderate (≥ 12 HDRS); *I* = 17.33; *C* = 15.34	HDRS IDAS PHQ-9 SCID-I SF-36	Baseline, week 2, week 4, week 6, week 8 (post-intervention)
DaCosta et al. (2009)	Canada	*N* = 88 (Ne = 46; Nc = 42)	NR (Me = 34.3; Mc = 32.7)	Elevated depressive symptoms (EPDS ≥ 10)	Mild to moderate (≥ 10 EPDS); 13.6 for both groups	EPDS HAM-D	Baseline, month 3 (post-intervention), month 6
Daley et al. (2015)	UK	*N* = 94 (Ne = 47; Nc = 47)	NR (Me = 31.7; Mc = 29.3)	Elevated depressive symptoms (EPDS ≥ 13) and CIS-R	Moderate (39% of participants had thoughts of self-harming, 18.1% with severe depression, 53.2% with a moderate-severe depression episode, 15.9% with a mild depression episode, and 12.8% with mixed anxiety and depressive disorder); *I* = 17.3; *C* = 17.5	EPDS SF-12	Baseline, month 6 (post-intervention), month 12
Daley et al. (2008)	UK	*N* = 38 (Ne = 20; Nc = 18)	Majority 21–40, (NR)	Clinical judgement or elevated depressive symptoms (EPDS ≥ 12)	Moderate to severe (most of the participants were taking medication for PND); *I* = 17.7; *C* = 19.2	EPDS	Baseline, week 12 (post-intervention)
Forsyth et al. (2017)	UK	*N* = 22 (Ne = 11; Nc = 11)	NR (Me = 25; Mc = 27)	Elevated depressive symptoms (EPDS ≥ 12) and SCID-PN	Moderate *I* = 17.6; *C* = 15.9	EPDS SCID-PN diagnosis	Baseline, week 12 (post-intervention), month 6
Haruna et al. (2013)	Japan	*N* = 101 (Ne = 50; Nc = 51)	NR (Me = 33.8; Mc = 33.7)	N/A	None *I* = 4.1; *C* = 5.9	EPDS SF-36v2	Baseline, month 2 (post-intervention)
Keller et al. (2014)	USA	*N* = 139 (Ne = 71; Nc = 68)	NR (*M* = 28.3)	N/A	Mild *I* = 8.21; *C* = 8.69	EPDS	Baseline, month 6 (post-intervention), month 12
LeCheminant et al. (2014)	USA	*N* = 60 (Ne = 30; Nc = 30)	NR (Me = 26.9; Mc = 25.9)	N/A	None *I* = 9.5; *C* = NR	CES-D	Baseline, month 2, month 4 (post-intervention)
Lewis et al. (2014)	USA	*N* = 130 (Ne = 66; Nc = 64)	NR (Me = 31.69; Mc = 31.39)	Personal history of depression or maternal history of depression but individuals with current depressive episodes were excluded	Mild (29% were taking antidepressant medication); *I* = 5.0; *C* = 5.0	EPDS PHQ-9 SCID-I	Baseline, month 6 (post-intervention)
Norman et al. (2010)	Australia	*N* = 161 (Ne = 80; Nc = 81)	17–41 (Me = 29.3; Mc = 30.1)	N/A	Mild *I* = 8.0; *C* = 6.75	EPDS	Baseline, week 8 (post-intervention), week 12

Table 2 (continued)

Studies	Country	*N* (Ne; Nc)	Age range in years, (mean)	Depression inclusion criteria	Baseline depression severity[a]; baseline depressive symptoms (mean)	Measures	Assessment time points
Robichaud (2009; unpublished thesis)	USA	*N* = 48 (Ne = 25; Nc = 23)	20–40 (Me = 31.1; Mc = 30.4)	N/A	Moderate to severe *I* = 19.76; *C* = 18.87	EPDS	Baseline, week 6 (post-intervention)
Saeedi et al. (2013)	Iran	*N* = 40 (Ne = 20; Nc = 20)	NR (Me 28.48; Mc 27.76)	Elevated depressive symptoms (EPDS ≥ 12)	Moderate to severe *I* = 19.14; *C* = 18.22	EPDS	Baseline, week 12 (post-intervention)
Shelton (2015; unpublished thesis)	USA	*N* = 6 (Ne = 3; Nc = 3)	NR (Me = 26.7; Mc = 25)	Elevated depressive symptoms (EPDS ≥ 7)	Mild *I* = 7.67; *C* = 9.33	EPDS	Baseline, week 6 (post-intervention)
Surkan et al. (2012)	USA	*N* = 679 (Ne = 337; Nc = 342)	18–44 (Me = 26.7; Mc = 26.3)	N/A	Moderate *I* = 14.3; *C* = 14.0	CES-D	Baseline, month 14 (post-intervention)
Thiruppathi et al. (2014)	India	*N* = 45 (Ne = 22; Nc = 23)	NR (Me = 26.3; Mc = 25.1)	N/A	Mild *I* = 7.95; *C* = 7.76	EPDS	Baseline, week 4 (post-intervention)
Yang and Chen (2018)	Taiwan	*N* = 140 (Ne = 70; Nc = 70)	NR (Me = 31.89; Mc = 32.45)	N/A	Mild *I* = 9.11; *C* = 8.45	EPDS	Baseline, week 4, week 12 (post-intervention)

CES-D Centre for Epidemiologic Studies Depression Scale, *DASS* Depression Anxiety Stress Scale, *EPDS* Edinburgh Postnatal Depression Scale, *GHQ12* 12-item General Health Questionnaire, *HAM-D/HDRS* Hamilton Depression Rating Scale, *IDAS* Inventory of Depression and Anxiety Symptoms, *M* mean age, *Mc* mean age of control group, *Me* mean age of experimental group, *N* sample size, *N/A* not applicable, *Nc* numbers in control group, *Ne* numbers in experimental group, *NR* not reported, *PHQ-9* Patient Health Questionnaire, *RCT* randomised controlled trial, *SCID-I* Structured Clinical Interview for DSM-IV Axis I Disorders, *SCID-PN* Structured Clinical Interview for DSM-IV (Perinatal Version), *SF-12* 12-Item Short-Form Health Survey, *SF-36* Medical Outcomes Study 36-Item Short-Form Health Survey

[a] Assessments were based on Cox et al. (1987), Kroenke et al. (2001), Radloff (1977), Zimmerman et al. (2013) and McCabe-Beane (2016)

Table 3 Characteristics of exercise and physical activity interventions

Studies	Type of intervention	Provider	Exercise type	Intervention duration (weeks or months)	Session duration (min); intensity	Session frequency per week	Intervention arm content and format (individual or group based)	Control arm content
Armstrong and Edwards (2003)	Targeted (indicated)	Supervised; NQ	Aerobic	12 weeks	30–40; moderate intensity	3 × exercise 1 × social support	Pram-walking and informal social support session Group-based	NI (circular walking test at baseline and post-intervention, plus an interim phone support session)
Armstrong and Edwards (2004)	Targeted (indicated)	Supervised and non-supervised; NQ	Aerobic	12 weeks	40; moderate intensity	2 × supervised exercise 1 × unsupervised exercise	Pram-walking Group-based	AC (non-structured, social support sessions (once per week)
Buttner et al. (2015)	Targeted (indicated)	Supervised and non-supervised; Q	Yoga	8 weeks	60 for supervised and 30 for unsupervised; NR	2 × supervised yoga 1 × unsupervised yoga (the minimum)	Sun salutations, balancing, twisting and relaxation poses Individual and group-based	WLC
DaCosta et al. (2009)	Targeted (indicated)	Supervised and non-supervised; Q	Aerobic exercise, plus coaching	12 weeks	90 for the first supervised and 30 for the 3 follow-up coaching sessions, 60–120/week unsupervised; moderate to high intensity	4 × supervised (within 12 weeks) Plus, individual weekly sessions	Stretching, strength and/or cardiovascular exercises, plus information and support elements Individual-based	UC
Daley et al. (2015)	Treatment	Non-supervised; N/A	Coaching (face-to-face exercise consultations and supportive telephone calls)	6 months	40–60 for the personalised consultations 15–20 for the telephone calls	2 × personalised exercise consultations (months 1 and 2) and telephone calls (months 3 and 4)	Promotion of physical exercise of moderate intensity on 3–5 days per week basis Individual-based	UC
Daley et al. (2008)	Targeted (indicated)	Non-supervised; N/A	Coaching	12 weeks	60 for the personalised consultations 10 for the telephone calls	2 × personalised exercise consultations over 12 weeks and 2 × follow-up support phone calls at weeks 3 and 9	Enhancing motivation and self-efficacy for undertaking moderate exercise on a weekly basis and preventing relapse Individual-based	UC
Forsyth et al. (2017)	Treatment	Supervised and non-supervised; NR	Coaching and aerobic exercise	12 weeks	60 for the personalised consultation and/or 150 of group-based or self-initiated exercise at moderate intensity (60 for each group-based session)	1 × personalised motivational consultation in 12 weeks (number of group-based and/or individual sessions per week is not reported)	Motivational and behaviour change coaching and pram-walking or facility based exercise Individual and group-based	UC
Haruna et al. (2013)	Universal	Supervised; Q	Aerobic	2 months	90	4 × supervised exercise	Aerobic and muscular stretching Group-based	NI
Keller et al. (2014)	Universal	Supervised; NQ	Coaching (social support) plus group walking	12 weeks	NR; moderate intensity	1 × supervised	Emotional, instrumental, appraisal and informational support plus group walking Group-based	AC (weekly telephone, informative sessions)
LeCheminant et al. (2014)	Universal	Supervised and non-supervised; Q	Resistance training for major muscle groups	18 weeks (4 months)	NR; mild to moderate intensity	2 × supervised and unsupervised	Leg extension, seated leg curl, leg press, biceps curl, shoulder press, chest press, seated row and abdominal curl-ups	AC (flexibility training)

Table 3 (continued)

Studies	Type of intervention	Provider	Exercise type	Intervention duration (weeks or months)	Session duration (min); intensity	Session frequency per week	Intervention arm content and format (individual or group based)	Control arm content
Lewis et al. (2014)	Targeted (selective)	Unsupervised; N/A	Coaching	6 months	NR; progressive intensity	1 × telephone coaching (month 1) 2 × telephone coaching per month (months 2 and 3) and 1 × telephone coaching per month (months 4, 5 and 6)	Individual-based Motivational strategies based on SCT and TTM Individual-based	AC wellness/support contact (11 phone-coaching sessions over 6 months)
Norman et al. (2010)	Universal	Supervised; Q	Aerobic, strengthening and coaching	8 weeks	60 min of supervised exercise sessions and 30 min of coaching session	1 × supervised exercise 1 × coaching	Cardiovascular and strength components Group-based	UC (education-only group)
Robichaud (2009)	Targeted (indicated)	Unsupervised; N/A	Aerobic	6 weeks	30 min walking, 45–60 min supervised initial session; NR	3 × unsupervised	Video/DVD-based exercise (walking) Individual-based	NI
Saeedi et al. (2013)	Targeted (indicated)	Supervised; NQ	Aerobic	12 weeks	45; NR	3 × supervised	Aerobic and stretching Group-based	NR
Shelton et al. (2015)	Universal	Unsupervised; N/A	Aerobic	6 weeks	30; moderate intensity	3 × unsupervised	Stroller-walking intervention at plus receiving education materials Individual-based	NI (receiving the education material only)
Surkan et al. (2012)	Universal	Unsupervised; N/A	Health promotion and coaching (home visits and telephone calls)	12 months	NR; N/A	5 × home visits (within 12 months) and 1 × phone call per month	Educational training, motivational interviewing and coaching includes objectives to perform 30 min of physical activity per day, at least 5 days per week Individual-based	UC (education training only)
Thiruppathi et al. (2014)	Universal	Supervised; Q	Aerobic, education and coaching	4 weeks	45; NR	1 × supervised exercise	Warm-up, cardiovascular intervals, body toning, core pelvic floor exercises, followed by cool down with stretching Individual-based	UC
Yang and Chen, 2018)	Universal	Unsupervised; N/A	DVD-based, Yoga, aerobic	3 months	15; progressive intensity	3 × unsupervised	Aerobic, muscle stretching and strengthening Individual-based	UC

N/A not applicable, *NI* non-intervention, *NQ* not qualified, *NR* not reported, *Q* qualified, *SCT* social cognitive theory, *TTM* transtheoretical model of exercise, *UC* usual care

participants ranged from 20 to 160, whilst one study had 679 participants. Apart from two included theses (Robichaud et al. 2009; Shelton 2015), the studies were published in peer-reviewed academic journals.

A targeted prevention approach was used in 10 studies, to target at-risk women with a history of depression or elevated depression symptoms (Armstrong and Edwards 2003; Armstrong and Edwards 2004; Buttner et al. 2015; DaCosta et al. 2009; Daley et al. 2008; Lewis et al. 2014; Robichaud et al. 2009; Saeedi 2013). A universal prevention approach (targeted at a whole population that has not been identified on the basis of individual risk) was tested in eight studies (Haruna et al. 2013; Keller et al. 2014; Norman et al. 2010; Shelton et al. 2015; Thiruppathi et al. 2014; Yang and Chen, 2018). Two studies tested a treatment approach for women with postpartum depression (Daley et al. 2015; Forsyth et al. 2017).

In six studies, participants' baseline depression symptoms were mild (Keller et al. 2014; Lewis et al. 2014; Norman et al. 2010; Shelton et al. 2015; Thiruppathi et al. 2014; Yang and Chen, 2018). In two studies, participants' symptoms were mild to moderate (Buttner et al. 2015; DaCosta et al.. 2009); in five studies, symptoms were moderate (Armstrong and Edwards 2003; Armstrong and Edwards 2004; Daley et al. 2015; Forsyth et al. 2017; Surkan et al. 2012), and in three studies, symptoms were moderate to severe (Daley et al. 2008; Robichaud et al. 2009; Saeedi, 2013).

Intervention and control conditions

Most studies compared the intervention arm to a NI or UC control condition, with four studies using an AC comparison (Armstrong and Edwards 2004; Keller et al. 2014; LeCheminant et al. 2014; Lewis et al. 2014). See Table 3 for an overview of intervention characteristics in each study.

In eight studies, the interventions tested were of aerobic and/or strengthening and/or muscle stretching content (Armstrong and Edwards 2004; Buttner et al. 2015; Haruna et al. 2013; LeCheminant et al. 2014; Robichaud 2009; Saaedi 2013; Shelton 2015; Yang and Chen, 2018). In four studies, the content was coaching and motivational health promotion techniques and no exercise (Daley et al. 2015; Daley et al. 2008; Lewis et al. 2014; Surkan et al. 2012). In six studies, the intervention followed a mixed approach of exercise and coaching/motivational promotion techniques (Armstrong and Edwards 2003; DaCosta et al. 2009; Forsyth et al. 2017; Keller et al. 2014; Norman et al. 2010; Thirrupathi et al. 2014).

The duration of 76% (13/17) interventions was up to 12 weeks, with four studies testing interventions for longer than 12 weeks (Daley et al. 2015; LeCheminant et al. 2014; Lewis et al. 2014; Surkan et al. 2012). The duration of the supervised delivered sessions ranged from 30 to 90 min, with most sessions delivered at moderate intensity. The frequency of the sessions delivered per week across the interventions ranged from 1 to 4.

Six studies were of supervised interventions (Armstrong and Edwards 2003; Haruna et al. 2013; Keller et al. 2014; Norman et al. 2010; Saeedi 2013; Thiruppathi et al. 2014), seven studies were of non-supervised interventions (Daley et al. 2015; Daley et al. 2008; Lewis et al. 2014; Robichaud 2009; Shelton 2015; Surkan et al. 2012; Yang and Chen, 2018) and five studies were of both supervised and non-supervised elements (Armstrong and Edwards 2004; Buttner et al. 2015; DaCosta et al. 2009; Forsyth et al. 2017; LeCheminant et al. 2014).

Of the supervised interventions, six were delivered by qualified service providers (Buttner et al., 2015; DaCosta et al. 2009; Haruna et al. 2013; LeCheminant et al. 2014; Norman et al. 2010; Thiruppathi et al. 2014), four were delivered by non-qualified service providers (Armstrong and Edwards 2003; Armstrong and Edwards 2004; Keller et al. 2014; Saeedi et al. 2013) and one did not report provider information (Forsyth et al. 2017). Table 3 presents an overview of intervention characteristics for each study.

Outcomes

Depression symptoms were assessed using the EPDS in most studies. Two studies used The Center for Epidemiological Studies-Depression (CES-D) (Surkan et al. 2012; LeCheminant et al. 2014) and one study used the Hamilton Rating Scale for Depression (HRSD) (Buttner et al. 2015). HRQoL was measured in three studies using the 36-Item Short-Form Health Survey (Buttner et al. 2015; Daley et al. 2015; Haruna et al. 2013), and anxiety symptoms were assessed in one study using the Inventory of Depression and Anxiety Symptoms (Buttner et al. 2015).

Quality assessment

Figure 2 presents the ratings for each item of the risk of bias assessment tool. Overall, most of the RCTs were of low to moderate quality. "Other risk of bias" was identified in multiple studies and was caused by (i) uncertainty about ITT analysis in five studies (Daley et al. 2008; Norman et al. 2010; Thiruppathi et al. 2014; Yang and Chen, 2018) and (ii) potential threat of unsuccessful randomisation in one study (Daley et al. 2015). "Unclear risk of bias" was identified in multiple studies caused by (i) insufficient details of the allocation concealment procedures and (ii) insufficient details regarding the sequence generation methods (five studies). There was poor reporting of the outcomes in two of the studies (Saeedi 2013; Thirrupathi et al. 2014) leading to a rating of high risk of bias. Given the nature of intervention and control conditions, a complete blinding procedure was impossible; however, given the outcome was self-report in most of the studies, they were generally rated as low-risk in the "blinding" sections of the risk of bias tool. Studies that reported an intention-to-treat analysis were rated as low risk of bias (Higgins et al. 2011).

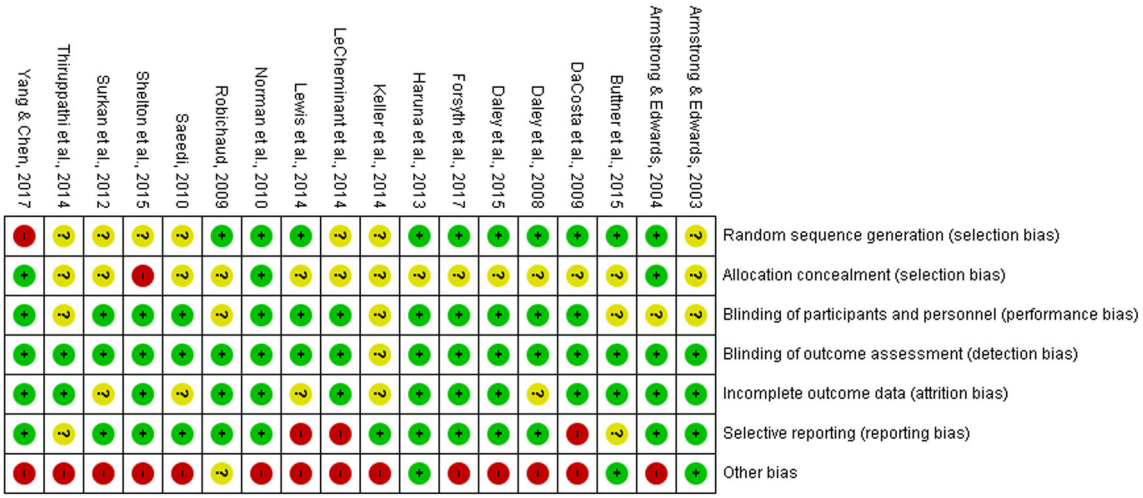

Fig. 2 Risk of bias assessment for each included study

Meta-analysis

A moderate, significant, SMD, favouring the intervention condition, was found for depressive symptoms (SMD = − 0.64, 95% CI = [− 0.96, − 0.33], $p < 0.001$; see Fig. 3 for forest plot including all studies and the bias-adjusted Hedge's g effect sizes). A non-significant SMD, favouring the intervention condition, was found for secondary outcomes: physical function (SMD = − 0.04, 95% CI = [− 0.33, 0.26], $p = 0.81$; and a non-significant SMD, favouring the control condition, was found for mental function (SMD = 0.27, 95% CI = [− 0.03, 0.56], $p = 0.07$). Due to the dearth of data, effect sizes for anxiety were not calculated.

Sensitivity analyses

Results of the sensitivity analyses showed a small, significant effect on depression, favouring the intervention condition (SMD = − 0.30, 95% CI = [− 0.45, − 0.15], $p < 0.001$) (Armstrong and Edwards 2004; Buttner et al. 2015; DaCosta et al. 2009; Daley et al. 2008; Daley et al. 2015; Forsyth et al. 2017; Haruna et al. 2013; Lewis et al. 2014; Norman et al. 2010; Robichaud 2009) (see Fig. 4). A post hoc sensitivity analysis compared the effectiveness of the exercise-based interventions after removing the two outlying studies (Saaedi 2013; Thirrupathi et al. 2014). This post hoc sensitivity analysis yielded small, significant, results (SMD = − 0.25, 95% CI = [− 0.39, − 0.11], $p = 0.0005$) (see Fig. 5).

Subgroup analyses

A comparison of the effectiveness of universal prevention interventions (Haruna et al. 2013; Keller et al. 2014; Norman et al. 2010; Shelton 2015; Surkan et al. 2012; Thiruppathi et al. 2014; Yang and Chen, 2018) vs. targeted prevention or treatment interventions (Armstrong and Edwards 2003; Armstrong and Edwards 2004; Buttner et al.

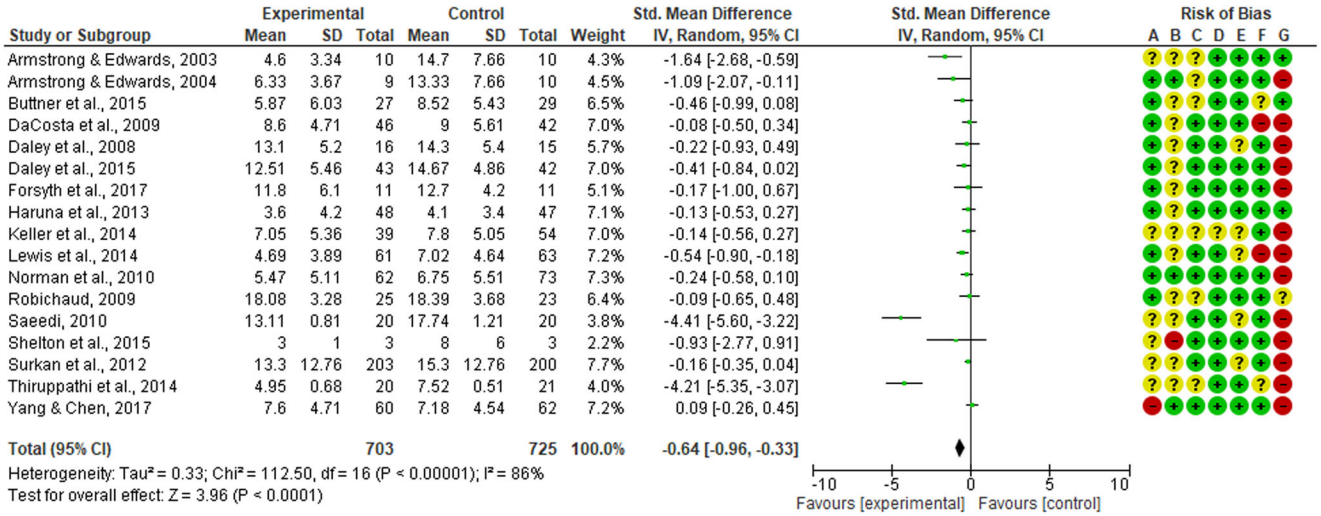

Fig. 3 Forest plot and effect size estimates for the effectiveness of exercise-based and physical activity interventions in reducing depressive symptoms

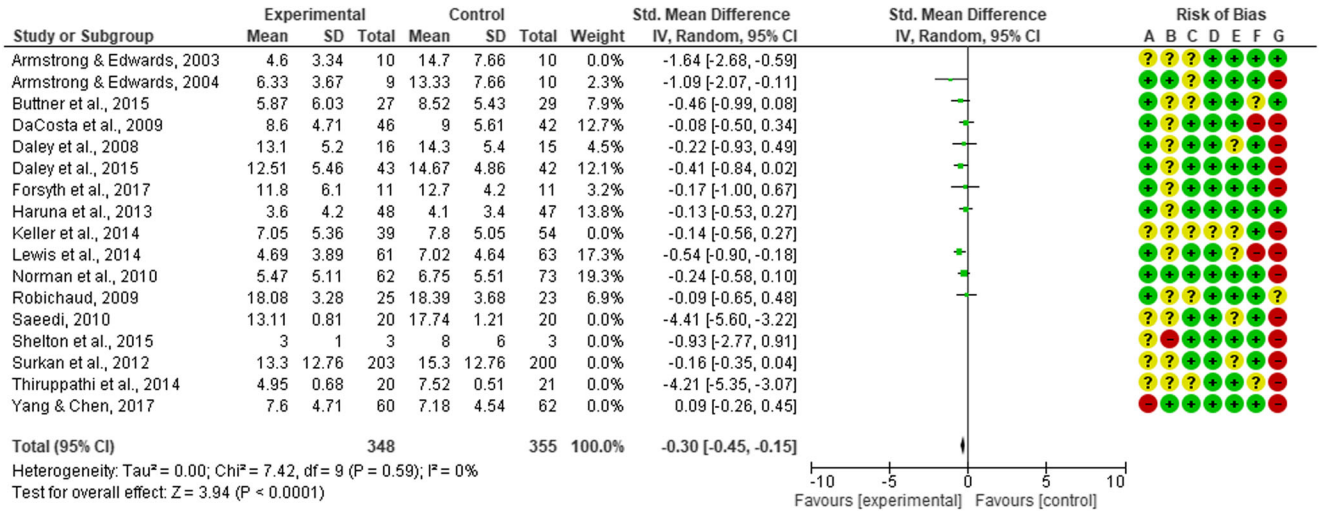

Fig. 4 Sensitivity analysis of studies rated as low risk of bias for random sequence generation

2015; DaCosta et al. 2009; Daley et al. 2008; Daley et al. 2015; Forsyth et al. 2017; Lewis et al. 2014; Robichaud 2009; Saeedi 2013) was conducted. Targeted prevention or treatment interventions yielded a greater effect size compared to universal prevention interventions (SMD = − 0.75, 95% CI = [− 1.22, − 0.28], $p = 0.002$ for the targeted interventions and SMD = − 0.52, 95% CI = [− 0.99, − 0.05], $p = 0.03$ for universal prevention interventions) (see Fig. 6).

A comparison of the effectiveness of interventions with an active exercise-oriented component (Armstrong and Edwards 2003; Armstrong and Edwards 2004; Buttner et al. 2015; DaCosta et al. 2009; Haruna et al. 2013; Norman et al. 2010; Robichaud 2009; Saeedi 2013; Shelton 2015; Thiruppathi et al. 2014) vs. those with coaching/motivational components (Daley et al. 2008; Daley et al. 2015; Forsyth et al. 2017; Keller et al. 2014; Lewis et al. 2014; Surkan et al. 2012; Yang and Chen, 2018) was conducted. Interventions with active exercise-oriented components yielded larger effects than those with coaching/motivational components (SMD = − 1.19, 95% CI = [− 1.84, − 0.53], $p = 0.0004$ for active exercise interventions and SMD = − 0.21, 95% CI = [− 0.37, − 0.05], $p = 0.009$ for coaching/motivational interventions (see Fig. 7).

A comparison of the effectiveness of the intervention arms against AC vs. the intervention arms against NI, UC and WLC was conducted. When tested against ACs (SMD = − 0.46, 95% CI = [− 0.86, − 0.05], $p = 0.03$), the exercise-based interventions yielded a smaller effect than those tested against NI, UC and WLC (SMD = − 0.70, 95% CI = [− 1.09, − 0.32], $p = 0.0003$) (see Fig. 8).

A comparison of interventions with long duration (12 weeks or more) vs. interventions with a shorter duration (fewer than 12 weeks) was conducted. Interventions with shorter duration (SMD = − 1.72, 95% CI = [− 3.05, − 0.39], $p = 0.01$) yielded a larger effect sizes than those of longer duration (SMD = − 0.52, 95% CI = [− 0.84, − 0.19], $p =$

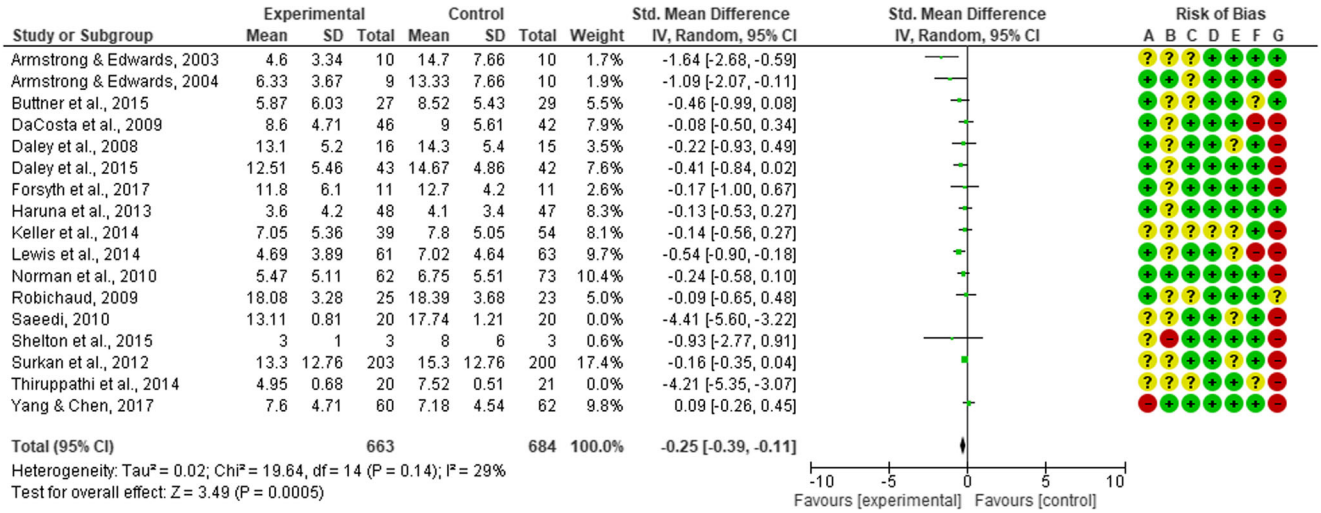

Fig. 5 Post hoc sensitivity analysis following removal of outliers

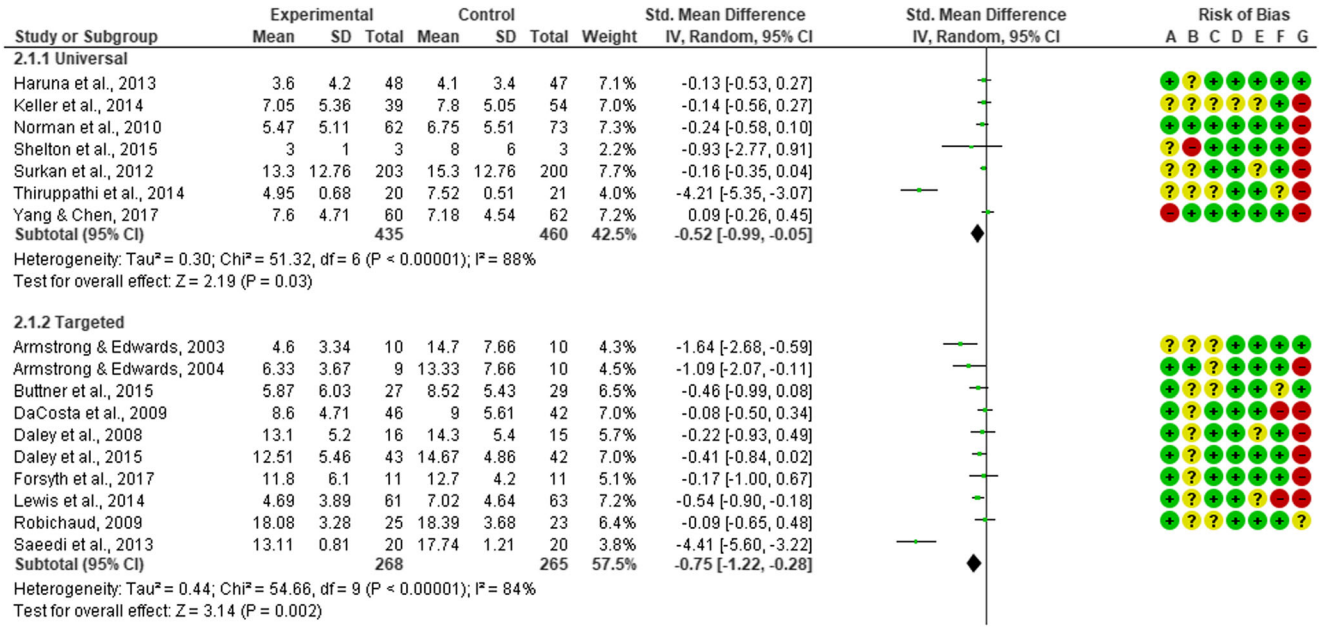

Fig. 6 Subgroup analysis of universal and targeted interventions

0.002) A meta-regression for the effect of duration on effect sizes of these interventions was performed with no significant results ($\beta = 0.07$, 95% CI = [− 0.11, 0.25], $p = 0.415$) (see Fig. 9).

Heterogeneity

Heterogeneity was high in the main analysis ($I^2 = 86\%$, $\text{Tau}^2 = 0.33$, $df = 16$, $p < 0.0001$) but was eliminated in the sensitivity analysis ($I^2 = 0\%$, $\text{Tau}^2 = 0$, $df = 9$, $p = 0.59$) where studies with no clear reporting of randomisation procedure were excluded.

Publication bias

Inspection of the funnel plot for the main analysis revealed extensive asymmetry (see Figs. 10 and 11 for the funnel plot and the contour-enhanced funnel plot), indicating potential threat for publication bias. An Egger's test was performed (Egger et al. 1997) for testing the funnel plot's asymmetry, indicating statistically significant results for small-study

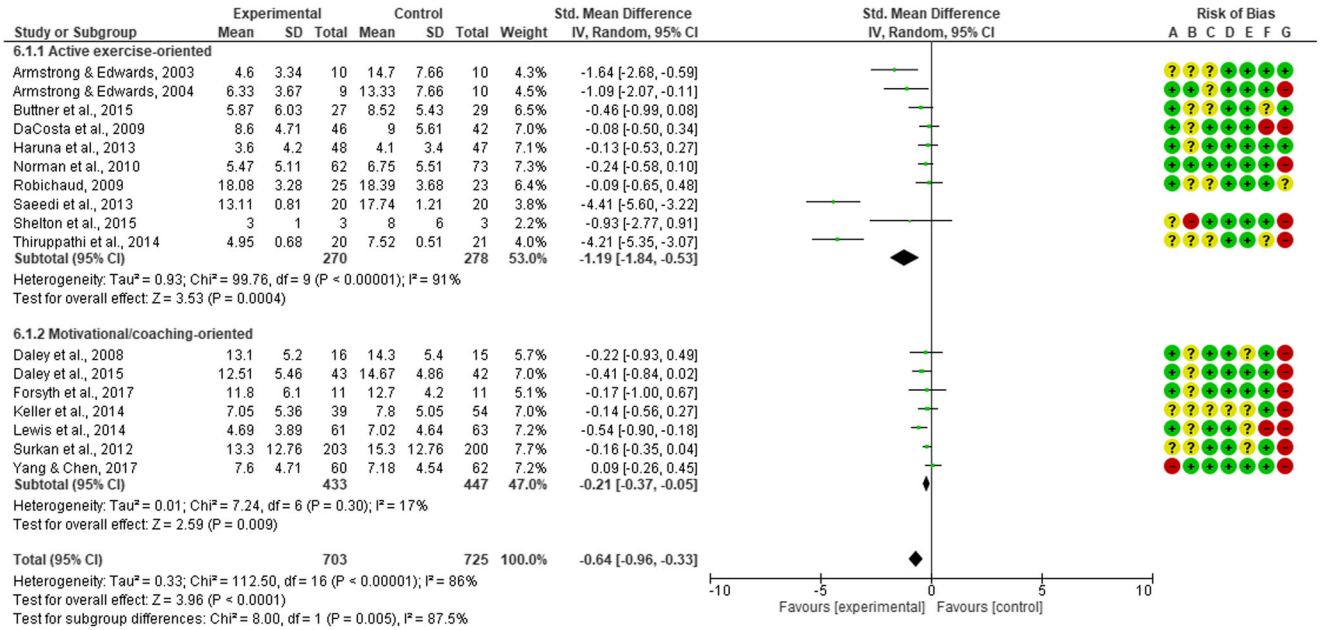

Fig. 7 Subgroup analysis of active exercise-orientated interventions vs. motivational/coaching-orientated interventions

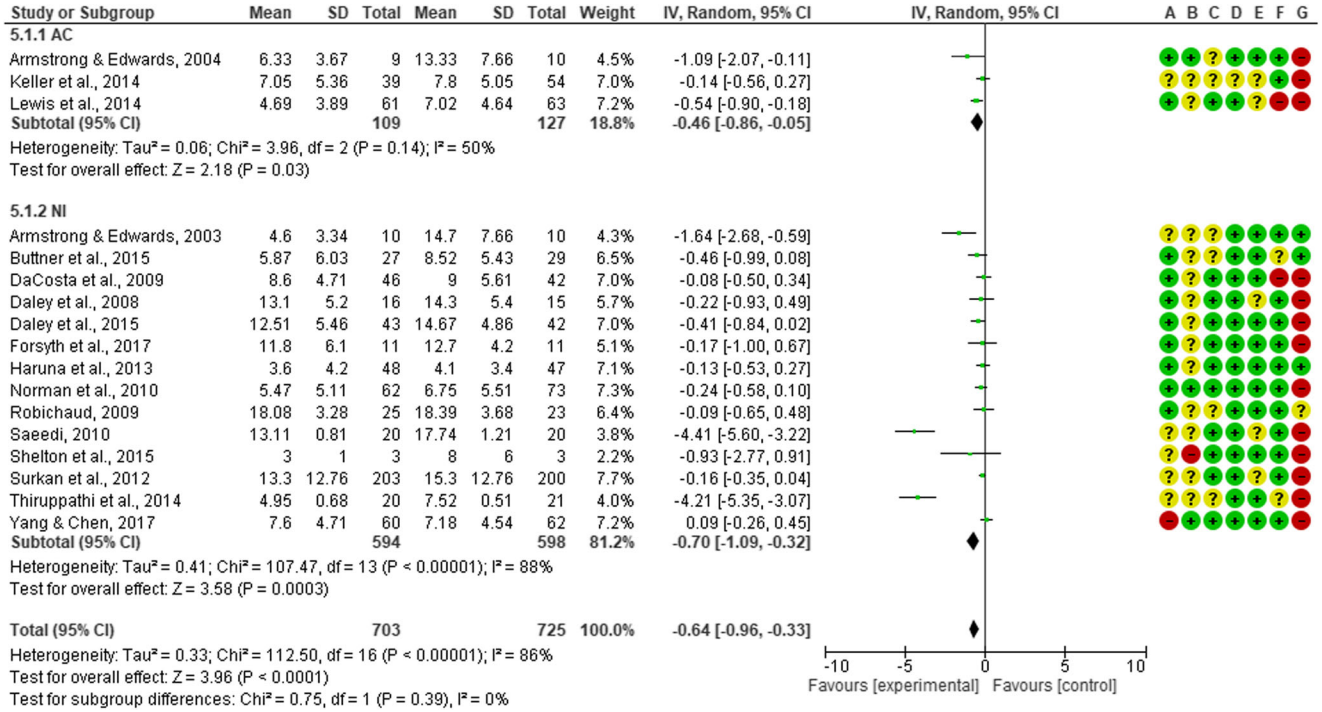

Fig. 8 Subgroup analysis of studies with active control conditions vs. other control conditions

effects ($\beta = -4.72$, 95% CI = [−5.44, −4.00], $p = 0.000$). However, after the two outlier studies were excluded, the Egger's test did not retain statistical significance ($\beta = -0.08$, 95% CI = [−0.29, 0.45], $p = 0.647$).

Rating the quality of evidence: the GRADE approach

Due to the dearth of data on secondary outcomes, the quality of evidence was assessed only for the primary outcome. Table 4 is a summary of findings (SoF) table that presents the comparison between exercise/physical activity-based interventions against all types of controls (AC, NI, UC, WL) in reducing depression symptoms. SMD is re-expressed as mean difference (MD) using a familiar instrument, the EPDS, in order to facilitate clinical interpretation (Ryan, Sontensso, and Hill 2016; Schunemann et al. 2008). To do so, a pooled standard deviation for EPDS scores was obtained from a cluster RCT (Morrell et al., 2009a, b) in order to transform SMD

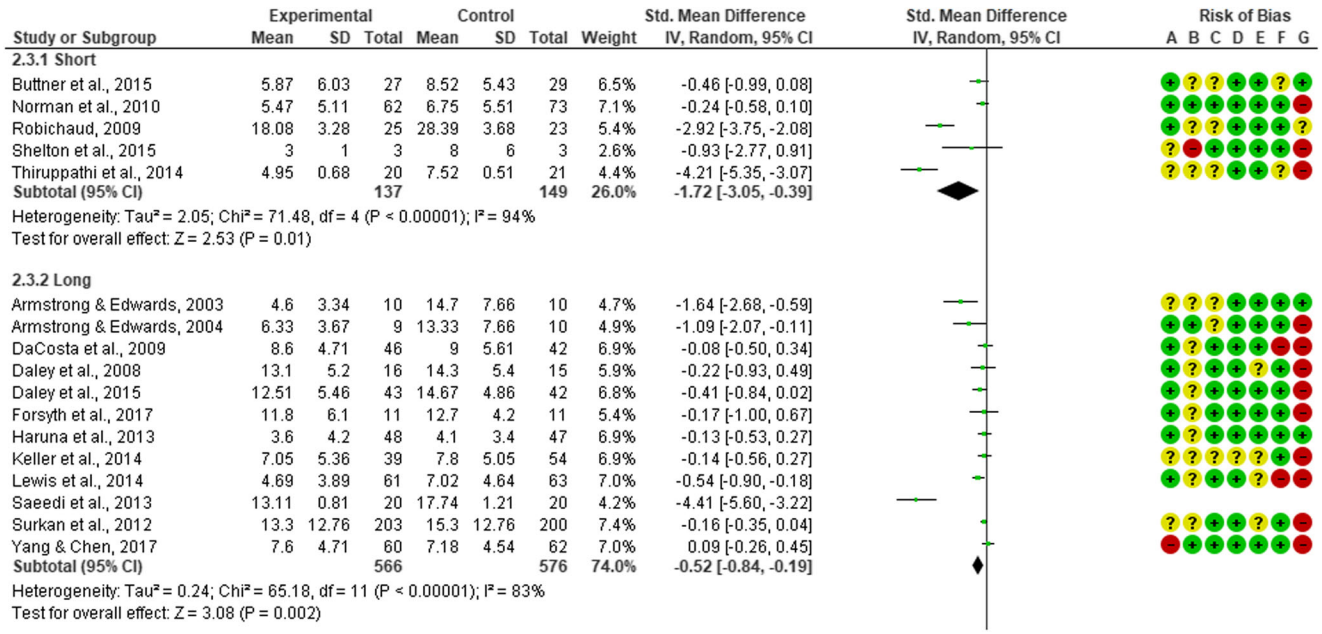

Fig. 9 Subgroup analysis of short and long duration interventions

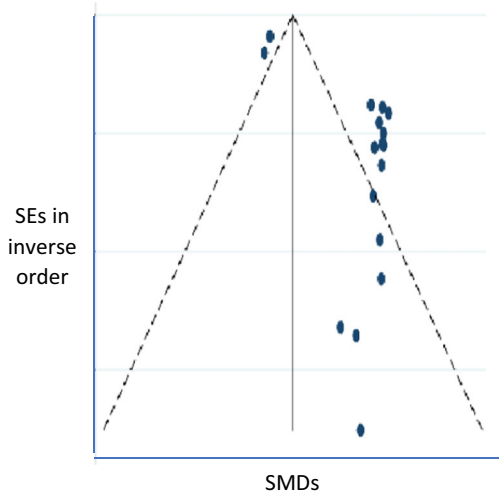

Fig. 10 Funnel plot with all the included studies

to MD. A small to moderate effect of exercise-based interventions to reduce depressive symptoms was found. We did not downgrade the quality of evidence regarding publication bias, given that the Egger test was non-significant after removing the two outlier studies (Saeedi 2013; Thirruppathi et al. 2014). However, since 76% (13/17) of the studies did not report a clear allocation concealment method, 41% (7/17) studies reported inadequate methods for sequence generation, and it was unclear whether some of the studies followed an ITT analysis, the quality of evidence was downgraded one level in the risk of bias section. In addition, the confidence intervals in most of the studies cross ± 0.50, leading to the downgrading of the quality of evidence regarding the imprecision of effects (Ryan and Hill 2016). The downgrading of the evidence was undertaken in accordance with established guidance (see Balshem et al. 2011). Consequently, the downgrading in two categories led to a low rating of the quality of evidence regarding the effectiveness of exercise-based interventions in reducing depression symptoms in postpartum women (Ryan and Hill 2016). Additionally, the transformation of SMD to MD, using a population-based SD for EPDS scores, highlighted that this mean difference does not signify a clinically significant difference (Matthey 2004). In summary, our confidence in the effect estimate for depression symptoms is limited: The true effect may be substantially different from the estimate of the effect.

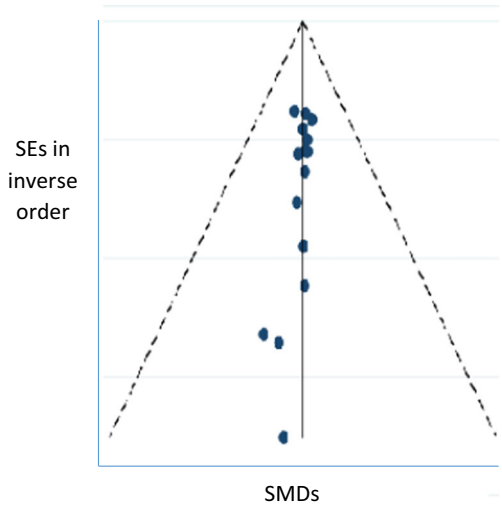

Fig. 11 Funnel plot of the included studies after removing the two outlier studies

Discussion

This meta-analysis found a statistically significant moderate treatment effect (SMD = − 0.64) of exercise over control conditions for depression symptoms in postpartum women up to 52 weeks after childbirth. Due to high levels of heterogeneity (I^2 = 86%), a sensitivity analysis was conducted excluding the studies with a high risk of bias. This analysis eliminated heterogeneity, however reduced the magnitude of effect to small (SMD = − 0.30), suggesting a consistent yet reduced effect of exercise for depression symptoms in postpartum women.

As the postpartum period can pose problems for managing weight in non-lactating women and for maintaining physical activity (Gaston and Cramp, 2011), the introduction of an exercise intervention is likely to have additional physical benefits alongside the effect of reducing symptoms of depression. Qualitative evidence suggests that additional benefits of exercise are improved confidence, body image and mood (Pritchett et al. 2017). Moreover, when lactating women are reluctant to take anti-depressant medication (Turner et al. 2008), exercise provides an acceptable alternative.

Subgroup analyses revealed that exercise-based interventions targeting at-risk women with a history of depression or elevated depression symptoms postpartum yielded increased treatment effects than universal preventive interventions. A similar finding has been reported previously in the postpartum population (McCurdy et al. 2017), and in young people (Carter et al. 2016), thus suggesting exercise interventions may be best applied as either a targeted preventive or treatment intervention. However, when exercise could be most efficacious, it is paradoxically when an individual might be less likely to undertake exercise due to the physical symptoms of depression (i.e. fatigue, diminished concentration, disturbed sleep and appetite) understandably adversely affecting motivation and activity levels. Consequently, future studies testing exercise for postpartum women with elevated depression symptoms need to focus on how to maximise appeal of the intervention and target motivation.

Importantly, the majority of the included studies did not assess anxiety symptoms despite the well evidenced comorbidity of anxiety and depression in the postpartum period (Falah-Hassani, Shiri and Denni 2016). Interestingly, this is not confined to exercise interventions as there is a reported

Table 4 GRADE table for assessing the quality of evidence

Certainty assessment							Summary of findings			Certainty	Importance[b]
No. of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	No. of pts. Exercise	Controls	Effect SMD MD[a] (95% CI)		
Outcome: depressive symptoms in postpartum women in 4 to 52 weeks Comparison: exercise-based and physical activity interventions vs. all types of controls (AC, NI, UC, WL) in reducing depressive symptoms in postpartum women receiving the allocated intervention within 4 and 52 weeks.											
17	RCT	Serious	Not serious	Not serious	Serious	None	703	725	− 0.64 (− 0.96, − 0.33)[c]	− 1.92 (− 2.88, − 0.99)	LOW	Non-clinically significant

[a] SMD have been re-expressed in MD using a familiar instrument (EPDS) in order to facilitate clinical interpretation (Ryan and Hill, 2016; Schunemann et al. 2008). A standard deviation for EPDS scores has been used from a large UK sample of women (Morrell et al. 2009)

[b] Clinical significance in EPDS change scores was based on Matthey

[c] I^2 = 86% for this effect estimate

general lack of research testing the efficacy/effectiveness of treatments for postnatal anxiety (Field 2018). As such, future studies should pay more attention to assessing and measuring symptoms of anxiety in pregnant and postnatal women with depression symptoms.

Strengths and limitations

This review has a number of strengths: (a) it is the first to include four RCTs of exercise for postpartum women that have not been previously included in qualitative and/or quantitative syntheses (Forsyth et al. 2017; LeCheminant et al. 2014; Thirruppathi et al. 2014; Yang and Chen, 2018); (b) it includes only RCTs, thus recommendations are based on the best quality available evidence; (c) all subgroup analyses undertaken included a sufficient number of studies, thus reducing the likelihood of making spurious recommendations; (d) it is the first in this area to follow the GRADE approach for rating the quality of evidence; (e) the reporting conforms to PRISMA guidance; and (f) the review has a prospectively registered protocol.

After careful inspection of the funnel plots, and without excluding the possibility of the publication bias, we assume that the poor methodological quality of smaller studies in this review has led to spuriously inflated effects (Sterne et al. 2008). The conclusions of the review are limited by the number and quality of the included studies. Although adequate numbers of participants were included to detect a difference in SMD as was found, the small number of studies limits the subgroup analysis possible. Moreover, due to the dearth of data on anxiety symptoms, no analysis was possible. In addition, the findings regarding the effects of exercise on HRQoL are limited, given that only two studies were included in the meta-analysis (Daley et al. 2015; Haruna et al. 2013). Finally, the overall low quality of the evidence limits the strength of the conclusions made.

Quality of evidence

The overall quality of evidence for exercise in depression symptoms in postpartum women is low, and our sensitivity analysis, which excluded studies at risk of selection bias, yielded a small treatment effect. Thus, the evidence does not currently support the large scale roll out of exercise interventions in treating and/or preventing depression symptoms in postpartum women.

Conclusion

Exercise is effective in reducing depression symptoms in postpartum women; however, the effect size is small to moderate and is based on mostly small, low-quality RCTs. The sensitivity analysis produced zero heterogeneity (I^2 = 0%) and

retained statistical significance, thus exercise as an intervention for postpartum depression symptom reduction certainly holds promise. Such an exercise intervention might be most effective for women with elevated symptoms of depression and delivered with increased focus on active engagement in supervised exercise sessions.

However, there is need for high-quality, sufficiently powered RCTs comparing exercise interventions against active controls. In addition, economic evaluations should be conducted in tandem with RCTs in order to assess the cost-effectiveness of exercise interventions for depression symptoms in postpartum women.

Acknowledgements This work was funded by the Nottingham City Clinical Commissioning Group's National Institute for Health Research Capability Funding

Compliance with ethical standards

Conflict of Interest Tim Carter, Anastasios Bastounis, Boliang Guo and C Jane Morrell declare that they have no conflict of interest.

References

Armstrong K, Edwards H (2003) The effects of exercise and social support on mothers reporting depressive symptoms: a pilot randomized controlled trial. Int J Mental Health Nurs 12(2):130–138

Armstrong K, Edwards H (2004) The effectiveness of a pram-walking exercise programme in reducing depressive symptomatology for postnatal women. Int J Nurs Pract 10(4):177–194

Austin MP, Hadzi-Pavlovic D, Priest SR, Reilly N, Wilhelm K, Saint K, Parker G (2010) Depressive and anxiety disorders in the postpartum period: how prevalent are they and can we improve their detection? Arch Womens Ment Health 13(5):395–401. https://doi.org/10.1007/s00737-010-0153-7

Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S et al (2011) GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 64(4):401–406

Bauer A, Parsonage M, Knapp M, Lemmi V, Bayo A (2014) Costs of perinatal mental health problems. London School of Economics and Political Science, London

Borenstein M, Hedges L, Higgins J, Rothstein H (2009) Introduction to meta-analysis. Wiley, Chichester

Buttner MM, Brock RL, O'Hara MW, Stuart S (2015) Efficacy of yoga for depressed postpartum women: a randomized controlled trial. Complement Ther Clin Pract 21(2):94–100. https://doi.org/10.1016/j.ctcp.2015.03.003

Campolong K, Jenkins S, Clark MM, Borowski K, Nelson N, Moore KM, Bobo WV (2017) The association of exercise during pregnancy with trimester-specific and postpartum quality of life and depressive symptoms in a cohort of healthy pregnant women. Arch Womens Ment Health. https://doi.org/10.1007/s00737-017-0783-0

Carter T, Morres ID, Meade O, Callaghan P (2016) The effect of exercise on depressive symptoms in adolescents: a systematic review and meta-analysis. J Am Acad Child Adolesc Psychiatry 55(7):580–590. https://doi.org/10.1016/j.jaac.2016.04.016

Carter T, Morrell J, Bastounis A, Guo B, Callaghan P (2017) Systematic review and meta-analysis of exercise interventions for preventing or treating postpartum depression. PROSPERO CRD42017068376. Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017068376.

Centre for Reviews and Dissemination (2009) Systematic reviews: CRD's guidance for undertaking reviews in health care. University of York, Centre for Reviews & Dissemination.

Claesson IM, Klein S, Sydsjo G, Josefsson A (2014) Physical activity and psychological well-being in obese pregnant and postpartum women attending a weight-gain restriction programme. Midwifery 30(1):11–16. https://doi.org/10.1016/j.midw.2012.11.006

Cooney GM, Dwan K, Greig CA, Lawlor DA, Rimer J, Waugh FR, Mead GE (2013) Exercise for depression. Cochrane Database Syst Rev 9:CD004366. https://doi.org/10.1002/14651858.CD004366.pub6

Da Costa D, Lowensteyn I, Abrahamowicz M, Ionescu-Ittu R, Dritsa M, Rippen N et al (2009) A randomized clinical trial of exercise to alleviate postpartum depressed mood. J Psychosom Obstet Gynaecol 30(3):191–200. https://doi.org/10.1080/01674820903212136

Daley AJ, Winter H, Grimmett C, McGuinness M, McManus R, MacArthur C (2008) Feasibility of an exercise intervention for women with postnatal depression: a pilot randomised controlled trial. Br J Gen Pract 58(548):178–183. https://doi.org/10.3399/bjgp08X277195

Daley AJ, Blamey RV, Jolly K, Roalfe AK, Turner KM et al (2015) A pragmatic randomized controlled trial to evaluate the effectiveness of a facilitated exercise intervention as a treatment for postnatal depression: the PAM-PeRS trial. Psychol Med 45(11):2413–2425. https://doi.org/10.1017/S0033291715000409

Egger M, Smith GD, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. BMJ 315(7109):629–634. https://doi.org/10.1136/bmj.315.7109.629

Falah-Hassani K, Shiri R, Dennis CL (2016) Prevalence and risk factors for comorbid postpartum depressive symptomatology and anxiety. J Affect Disord. 198:142–147

Farias-Antunez S, Xavier MO, Santos IS (2017) Effect of maternal postpartum depression on offspring's growth. J Affect Disord 228:143–152. https://doi.org/10.1016/j.jad.2017.12.013

Fazzi C, Saunders DH, Linton K, Norman JE, Reynolds RM (2017) Sedentary behaviours during pregnancy: a systematic review. Int J Behav Nutr Phys Act 14(1):32. https://doi.org/10.1186/s12966-017-0485-z

Field T (2018) Postnatal anxiety prevalence, predictors and effects on development: a narrative review. Infant Behav Dev. 51:24–32

Forsyth J, Boath E, Henshaw C, Brown H (2017) Exercise as an adjunct treatment for postpartum depression for women living in an inner city—a pilot study. Health Care Women Int 38(6):635–639. https://doi.org/10.1080/07399332.2017.1295049

Gaston A, Cramp A (2011) Exercise during pregnancy: a review of patterns and determinants. J Sci Med Sport 14(4):299–305. https://doi.org/10.1016/j.jsams.2011.02.006

Guyatt GH, Oxman AD, Vist GE Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 336(7650):924–926. https://doi.org/10.1136/bmj.39489.470347.AD

Haruna M, Watanabe E, Matsuzaki M, Ota E, Shiraishi M, Murayama R, Yeo S (2013) The effects of an exercise program on health-related quality of life in postpartum mothers: a randomized controlled trial. Health 05(03):432–439. https://doi.org/10.4236/health.2013.53058

Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Sterne JAC (2011) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 343. https://doi.org/10.1136/bmj.d5928

Keller C, Ainsworth B, Records K, Todd M, Belyea M, Vega-López S et al (2014) A comparison of a social support physical activity

intervention in weight management among post-partum Latinas. BMC Public Health 14(1):971. https://doi.org/10.1186/1471-2458-14-971

LeCheminant JD, Hinman T, Pratt KB, Earl N, Bailey BW, Thackeray R, Tucker LA (2014) Effect of resistance training on body composition, self-efficacy, depression, and activity in postpartum women. Scand J Med Sci Sports. 24(2):414–421

Lewis BA, Gjerdingen DK, Avery MD, Sirard JR, Guo H, Schuver K, Marcus BH (2014) A randomized trial examining a physical activity intervention for the prevention of postpartum depression: the healthy mom trial. Ment Health Phys Act 7(1):42–49. https://doi.org/10.1016/j.mhpa.2013.11.002

Matthey S (2004) Calculating clinically significant change in postnatal depression studies using the Edinburgh Postnatal Depression Scale. Journal of Affective Disorders 78(3):269–272. https://doi.org/10.1016/s0165-0327(02)00313-0

McCurdy AP, Boule NG, Sivak A, Davenport MH (2017) Effects of exercise on mild-to-moderate depressive symptoms in the postpartum period: a meta-analysis. Obstet Gynecol 129(6):1087–1097. https://doi.org/10.1097/aog.0000000000002053

Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 339. https://doi.org/10.1136/bmj.b2535

Morrell CJ, Slade P, Warner R, Paley G, Dixon S, Walters SJ, Brugha T, Barkham M, Parry G, Nicholl J (2009a) Clinical effectiveness of health visitor training in psychologically informed approaches for depression in postnatal women—pragmatic cluster randomised trial in primary care. BMJ 338:a3045

Morrell CJ, Warner R, Slade P, Dixon S, Walters S, Paley G, Brugha T (2009b) Psychological interventions for postnatal depression: cluster randomised trial and economic evaluation. The PoNDER trial. Health Technol Assess 13(30):1–153

Morrell CJ, Sutcliffe P, Booth A, Stevens J, Scope A, Stevenson M et al (2016) A systematic review, evidence synthesis and meta-analysis of quantitative and qualitative studies evaluating the clinical effectiveness, the cost-effectiveness, safety and acceptability of interventions to prevent postnatal depression. Health Technol Assess 20(37)

National Institute for Health and Care Excellence (2014) Antenatal and postnatal mental health: clinical management and service guidance. Clinical guideline [CG192]. Retrieved from https://www.nice.org.uk/guidance/cg192

Norman E, Sherburn M, Osborne RH, Galea MP (2010) An exercise and education program improves well-being of new mothers: a randomized controlled trial. Phys Ther 90(3):348–355. https://doi.org/10.2522/ptj.20090139

Pope CJ, Xie B, Sharma V, Campbell MK (2013) A prospective study of thoughts of self-harm and suicidal ideation during the postpartum period in women with mood disorders. Arch Womens Ment Health 16(6):483–488. https://doi.org/10.1007/s00737-013-0370-y

Poyatos-Leon R, Garcia-Hermoso A, Sanabria-Martinez G, Alvarez-Bueno C, Cavero-Redondo I, Martinez-Vizcaino V (2017) Effects of exercise-based interventions on postpartum depression: a meta-analysis of randomized controlled trials. Birth 44(3):200–208. https://doi.org/10.1111/birt.12294

Pritchett RV, Daley AJ, Jolly K, (2017) Does aerobic exercise reduce postpartum depressive symptoms? a systematic review and meta-analysis. British Journal of General Practice 67 (663):e684-e691

Putnam KT, Wilcox M, Robertson-Blackmore E, Sharkey K, Bergink V et al (2017) Clinical phenotypes of perinatal depression and time of symptom onset: analysis of data from an international consortium. Lancet Psychiatry. 4(6):477–485

Robichaud KP (2009) The effects of an exercise intervention on the psychological well-being of postpartum women. Diss Abstr Int: Section B: Sci Eng 69(7-B):4124

Ryan R, Hill S (2016). How to GRADE the quality of the evidence. Cochrane Consumers and Communication Group. Retrieved from http://cccrg.cochrane.org/author-resources.

Saeedi S (2013) Effect of exercise program on symptoms of postpartum depression. Iran J Obstet, Gynecol Infertility 15(34):26–31

Saligheh M, Hackett D, Boyce P, Cobley S (2017) Can exercise or physical activity help improve postnatal depression and weight loss? A systematic review. Arch Womens Ment Health 20(5):595–611. https://doi.org/10.1007/s00737-017-0750-9

Schunemann H, Oxman A, Higgins J, Vist, G, Glasziou P, Guyatt G, & on behalf of the Cochrane Applicability and Recommendations Methods Group and the Cochrane Statistical Methods Group. (2008). Presenting results and 'Summary of findings' tables. In J. Higgins & S. Green (Eds.), Cochrane handbook for systematic reviews of interventions. Cochrane book series. Chichester: Wiley-Blackwell.

Shelton S (2015) Postpartum depressive symptoms: a study of influencing factors and an intervention for improvement. Dissertation, Georgia State University.

StataCorp (2015) Stata statistical software: release 14. College Station, TX: StataCorp LP.

Sterne J, Egger M, Moher D (2008) Chapter 10: addressing reporting biases. In: Higgins J, Green S (eds) Cochrane handbook for systematic reviews of interventions, Cochrane book series. Wiley-Blackwell, Chichester, UK

Surkan PJ, Gottlieb BR, McCormick MC, Hunt A, Peterson KE (2012) Impact of a health promotion intervention on maternal depressive symptoms at 15 months postpartum. Matern Child Health J 16(1):139–148. https://doi.org/10.1007/s10995-010-0729-x

Teychenne M, York R (2013) Physical activity, sedentary behavior, and postnatal depressive symptoms: a review. Am J Prev Med 45(2):217–227. https://doi.org/10.1016/j.amepre.2013.04.004

Thiruppathi A, Prasana B, Mastaniah E, Vamsidhar N, Himabindu P (2014) A structured physical activity and health care education beats postpartum depression for primipara mothers: a pilot randomized controlled trial. Int J Physiother 1(3):144. https://doi.org/10.15621/ijphy/2014/v1i3/53468

Turner KM, Sharp D, Folkes L, Chew-Graham C, (2008) Women's views and experiences of antidepressants as a treatment for postnatal depression: a qualitative study. Family Practice 25 (6):450-455

Wisner KL, Sit DK, McShea MC, Rizzo DM et al (2013) Onset timing, thoughts of self-harm, and diagnoses in postpartum women with screen-positive depression findings. JAMA Psychiatry 70(5):490–498. https://doi.org/10.1001/jamapsychiatry.2013.87

World Health Organisation (WHO) 2017 Maternal mental health. WHO http://www.who.int/mental_health/maternal-child/maternal_mental_health/en/. Accessed 02 February 2018.

Yang CL, Chen CH (2018) Effectiveness of aerobic gymnastic exercise on stress, fatigue, and sleep quality during postpartum: a pilot randomized controlled trial. Int J Nurs Stud 77:1–7. https://doi.org/10.1016/j.ijnurstu.2017.09.009