Impact of prenatal exercise on both prenatal and postnatal anxiety and depressive symptoms: a systematic review and meta-analysis


ABSTRACT

Objective To examine the influence of prenatal exercise on depression and anxiety during pregnancy and the postpartum period.

Design Systematic review with random effects meta-analysis and meta-regression.

Data sources Online databases were searched up to 6 January 2017.

Study eligibility criteria Studies of all designs were included (except case studies) if they were published in English, Spanish or French and contained information on the Population (pregnant women without contraindication to exercise), Intervention (subjective or objective measures of frequency, intensity, duration, volume or type of exercise), Comparator (no exercise or different frequency, intensity, duration, volume and type of exercise) and Outcome (prenatal or postnatal depression or anxiety).

Results A total of 52 studies (n=131 406) were included. ‘Moderate’ quality evidence from randomised controlled trials (RCTs) revealed that exercise-only interventions, but not exercise+cointerventions, reduced the severity of prenatal depressive symptoms (13 RCTs, n=1076; standardised mean difference: −0.38, 95% CI −0.51 to −0.25, I²=10%) and the odds of prenatal depression by 67% (5 RCTs, n=683; OR: 0.33, 95% CI 0.21 to 0.53, I²=0%) compared with no exercise. Prenatal exercise did not alter the odds of postpartum depression or the severity of depressive symptoms, nor anxiety or anxiety symptoms during or following pregnancy. To achieve at least a moderate effect size in the reduction of the severity of prenatal depressive symptoms, pregnant women needed to accumulate at least 644 MET-min/week of exercise (eg, 150 min of moderate intensity exercise, such as brisk walking, water aerobics, stationary cycling, resistance training).

Summary/Conclusions Prenatal exercise reduced the odds and severity of prenatal depression.

INTRODUCTION

In 2017, the WHO identified that depression is the leading cause of disability worldwide. Pregnant and postpartum women are particularly vulnerable to declining mental health as a result of elevated stress and lack of social support. Depression affects approximately 13% and anxiety affects up to 39% of pregnant and postpartum women; however, it is generally accepted that both conditions are underdiagnosed and undertreated during the perinatal period. These conditions have immediate and persisting adverse impacts on both mother and child.

For mothers, depression and/or anxiety during pregnancy strongly predicts postpartum depression and anxiety; depression and anxiety are associated with reduced self-care, preterm labour, low birth weight as well as compromised caregiving and bonding with the infant. Maternal depression and anxiety have also been linked with delays in offspring development (cognitive, emotional and social development). Standard treatments for depression and anxiety include pharmacological and psychological intervention but the long-term impact of some medications on the fetus has not been established and psychotherapy can be costly and difficult to access. The consequences of leaving these disorders untreated can be severe.

In non-pregnant women, exercise may be as effective as antidepressants and psychotherapy for treating mild-to-moderate depression (ie, a moderate effect size). A recent review demonstrated that light to moderate intensity aerobic exercise initiated in the postpartum period improves mild to moderate depressive symptoms and increases the likelihood of resolving mild to moderate depression in the postpartum period compared with standard care. It has not been established if prenatal exercise reduces the prevalence and/or severity of depression and anxiety during pregnancy and the postpartum period.

Current national and international guidelines recommend that women without contraindications to exercise (eg, preeclampsia, incompetent cervix, premature labour) be physically active throughout pregnancy. This systematic review and meta-analysis forms part of a series of reviews, which will form the evidence base for the development of the 2019 Canadian guideline for physical activity throughout pregnancy (herein referred to as Guideline). The purpose of this paper is to present the results of a systematic review and meta-analysis of the relationships between prenatal exercise, and (1) depression and anxiety during pregnancy as well as (2) any persisting effect into the postpartum period.
METHODS
In October 2015, the Guidelines Consensus Panel was formed to identify outcomes for the Guideline update. The panel included researchers, methodological experts, a fitness professional, a public health representative (Middlesex-London Health Unit) and representatives from the Society for Obstetricians and Gynecologists of Canada, Canadian Society for Exercise Physiology (CSEP), The College of Family Physicians of Canada, Canadian Association of Midwives and Canadian Academy of Sport and Exercise Medicine, Exercise is Medicine Canada. During this meeting, 20 ‘critical’ and 17 ‘important’ outcomes related to prenatal exercise and maternal/fetal health were selected. Maternal mental health was rated as ‘critical’ outcome. This systematic review and meta-analysis was conducted in accordance with the PRISMA guidelines and the checklist was completed.25

Protocol and registration
Two systematic reviews examining the impact of prenatal exercise on fetal and maternal health outcomes were registered with PROSPERO, the international prospective register of systematic reviews (fetal health: Registration no. CRD42016029869; Available from: https://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016029869;maternal health: Registration no. CRD42016032376; Available from: https://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016032376). Because the relationships between prenatal exercise and maternal/fetal health outcomes are examined in studies related to both fetal and maternal health, records retrieved from both of these reviews were considered for inclusion in the present systematic review.

Eligibility criteria
This study was guided by the participants, interventions, comparisons, outcomes and study design (PICOS) framework.25

POPULATION
The population of interest included in this review was pregnant women without contraindication to exercise (according to the CSEP and American College of Obstetricians and Gynecologists (ACOG) guidelines).21 26 Absolute contraindications to exercise were defined as: ruptured membranes, premature labour, persistent second or third trimester bleeding, placenta praevia, preeclampsia, gestational hypertension, incompetent cervix, intrauterine growth restriction, high order pregnancy, uncontrolled type 1 diabetes, hypertension or thyroid disease or other serious cardiovascular, respiratory or systemic disorders. Relative contraindications to exercise were defined as: a history of spontaneous abortion, premature labour mild/moderate cardiovascular or respiratory disease, anaemia or iron deficiency, malnutrition or eating disorder, twin pregnancy after 28 weeks or other significant medical conditions.21 22 26

INTERVENTION (EXPOSURE)
The intervention/exposure was subjective or objective measures of frequency, intensity, duration, volume or type of exercise. Although exercise is a subtype of physical activity, for the purpose of this review, we used the terms interchangeably. Exercise and physical activity were defined as any bodily movement generated by skeletal muscles that resulted in energy expenditure above resting levels.27 Acute (ie, a single exercise session) or habitual (ie, usual activity) prenatal exercise and interventions including exercise alone (termed ‘exercise-only’ interventions) or in combination with other interventions (such as diet; termed ‘exercise+cointerventions’) were considered. Studies were excluded if exercise was performed after the beginning of labour.

COMPARISON
Eligible comparators were: various frequencies, intensities, durations, volumes and types of exercise; various durations of intervention or exercise or exercise in different trimesters.

OUTCOME
Eligible outcomes were diagnosis of depression or anxiety (either by exceeding a specific threshold via questionnaire or diagnosis by a qualified professional) and severity of symptoms of depression or anxiety.

STUDY DESIGN
Primary studies of any design were eligible, except case studies (n=1). Narrative or systematic reviews and meta-analyses were excluded.

Information sources
A comprehensive search was created and run by a research librarian (LGS) in the following databases: MEDLINE, EMBASE, PsycINFO, Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials, Scopus and Web of Science Core Collection, CINAHL. Plus with Full-text, Child Development & Adolescent Studies, ERIC, Sport Discus, ClinicalTrials.gov and the Trip Database up to 6 January 2017 (see online supplementary file for complete search strategies).

Study selection and data extraction
Titles and abstracts of all retrieved articles were independently screened by two reviewers. Abstracts that were judged to have met the initial screening criteria by at least one reviewer were retrieved as full-text articles. Articles were reviewed for relevant PICOS information by at least one person. If it was deemed that articles did not meet the inclusion criteria, they were reviewed by MHD and/or SMR prior to exclusion. If agreement could not be reached by discussion, study characteristics related to eligibility were presented to the Guidelines Steering Committee (MHD, MFM, SMR, CG, VP, AJG and NB) and final decisions were made by consensus. Studies that were selected were imported into DistillerSR (Evidence Partners, Ottawa, Ontario, Canada) for data extraction. At this point, studies from the maternal and fetal reviews that were included were de-duplicated against one another in DistillerSR and were considered as one review from this point forward.

Data extraction tables were created in DistillerSR in consultation with methodological experts and the Guidelines Steering Committee. Data were extracted by one person; a content expert (MHD, MFM or SMR) independently verified the extracted data. Reviewers were not blinded to study authors. For each single study, the most recent or complete version of the publication was selected as the ‘parent’ paper; however, relevant data from all publications related to each unique study were extracted. Study characteristics (ie, year, study design, country) and population characteristics (eg, number of participants, age, pre-pregnancy body mass index (BMI), previous physical activity levels, parity and pregnancy complications including preeclampsia, gestational hypertension and gestational diabetes), intervention/exposure (actual and/or prescribed exercise frequency, intensity, duration and type, duration of the intervention, measure of physical activity) and outcomes (depression or anxiety diagnosis and/or
Quality of evidence assessment
The Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework was used to assess the quality of evidence across studies for each study design and health outcome. The risk of bias for all included studies was independently assessed by two reviewers. The risk of bias in randomised controlled trials (RCTs) and intervention studies was evaluated following the Cochrane Handbook;28 risk of bias in observational studies was assessed using the characteristics recommended by Guyatt et al.29 All studies were examined for potential sources of bias, ie, selection bias, reporting bias, performance bias, detection bias, attrition bias and ‘other’ sources of bias.

Risk of bias across studies was rated as ‘serious’ when studies having the greatest influence on the pooled result (assessed using weight (%) given in forest plots or sample size in studies that were narratively synthesised) presented ‘high’ risk of bias. The greatest influence on the pooled result was determined as follows: the studies that had the greatest individual per cent contribution in the meta-analyses, when taken together, contributed to >50% of the weight of the pooled estimate. Serious risk of bias was considered when sample size of studies that were narratively synthesised was similar to the total sample size of studies contributing to >50% of the weight of the pooled estimate in the meta-analyses. Performance bias was rated as ‘high’ when <60% of participants performed 100% of prescribed exercise sessions or attended 100% of counselling sessions (defined as low compliance) or when compliance to the intervention was not reported. Attrition bias was rated as ‘high’ when >10% of data were missing at the end of the study and intention-to-treat analysis was not used. Given the nature of exercise interventions, it is not possible to blind participants to group allocation. Therefore, if the only source of bias was related to the blinding of allocation, the risk of bias was rated as ‘low’.

The risk of bias, indirectness, inconsistency, imprecision or risk of publication bias was assessed across studies for each study design and health outcome. Evidence from RCTs began with a ‘high’ quality of evidence rating and was graded down if any quality violations occurred. Evidence from all non-randomised intervention and observational studies began with a ‘low’ quality rating and, if there was no cause to downgrade, was upgraded if applicable according to the GRADE criteria (eg, large magnitude of effect, evidence of dose-response).28

Indirectness was considered serious when exercise-only interventions and exercise+cointerventions were combined for analysis or when the effect of exercise+cointervention on an outcome was examined. Exercise-only interventions could include standard care. Inconsistency was considered serious when heterogeneity was high (I²≥50%) or when only one study was assessed (I² unavailable). Imprecision was considered serious when the 95%CI crossed the line of no effect and was wide, such that interpretation of the data would be different if the true effect were at one end of the CI or the other. When only one study was included, imprecision was not considered serious because inconsistency was already considered serious for this reason. Finally, in order to assess publication bias, funnel plots were created if at least 10 studies were included in the forest plot (see online supplementary figure 2). If there were fewer than 10 studies, publication bias was deemed non-estimable and not rated down. Due to time constraints and feasibility, one reviewer evaluated the quality of the evidence across each health outcome using the protocol and a second person reviewed the GRADE tables as a quality control measure. Quality of evidence assessment is presented in online supplementary tables 2–5.

Statistical analysis
Statistical analyses were conducted using Review Manager V.5.3. (Cochrane Collaboration, Copenhagen, Denmark). ORs were calculated for all dichotomous outcomes. Inverse-variance weighting was applied to obtain OR using a random effects model. Standardised mean differences (SMD) were calculated when different measurement tools or scales were used for a single outcome. SMD effect sizes were calculated using Hedges’ g. An effect size of 0.2, 0.4 and 0.8 was considered small, moderate and large, respectively. Significance was defined as p<0.05. Meta-analyses were performed separately by study design. A staged approach was used to determine if there was sufficient evidence from RCTs for each outcome to inform the Guideline or if it was necessary to consider other study designs. If fewer than 2000 women were included in the exercise-only RCT meta-analysis, the impact of prenatal exercise on the specific outcome was examined further using observational evidence (non-randomised intervention, cohort, cross-sectional and case-control studies). For RCTs and non-randomised interventions, sensitivity analyses were performed to evaluate whether the effects were different when examining relationships between exercise-only interventions versus exercise+cointerventions and depression and anxiety. When possible, the following a priori subgroup analyses were conducted for exercise-only interventions: (1) women diagnosed with diabetes (gestational, type 1 or type 2) compared with women without diabetes (named ‘general population’); (2) samples of women with overweight or obesity (mean BMI>25.0 kg/m²) prior to pregnancy compared with samples of women who were of various BMI (mean BMI<25.0 kg/m² but possibly with some individuals with BMI>25.0 kg/m²; named ‘general population’); (3) women>35 years of age compared with women<35 years of age; (4) women who were previously inactive compared with those who were previously active (as defined by individual study authors). If a study did not provide sufficient detail to allow it to be grouped into the a priori subgroups, then a third group called ‘unspecified’ was created. Tests for subgroup differences were conducted, with statistical significance set at p<0.05. If statistically significant differences were found subgroup differences were interpreted. Finally, a priori subgroup analyses were also conducted for exercise-only RCTs with critical outcomes to identify whether a specific type of exercise was associated with greater benefit compared with aerobic or resistance training alone. Posthoc subgroup analyses were conducted comparing women with and without depression at baseline, women who exercised in supervised versus unsupervised settings and comparing effects in which the Edinburgh Postnatal Depression Scale (EPDS) versus Center for Epidemiologic Studies Depression Scale (CES-D) tools were used to assess possible diagnosis of depression and symptom severity. The I² was calculated to indicate the per cent of total variability that was attributable to between-study heterogeneity. In studies where there were no observed events in the intervention or control groups, data were entered into forest plots, but were considered ‘not estimable’ and excluded from the pooled analysis as per the recommendation in the Cochrane Handbook.30 Dose-response meta-regression31 32 was carried out by weighted no-intercept regression of log OR with a random effects model. The metafor package in R33 V.3.4.1. Models did not include an intercept term since the log OR is assumed to be zero when the exercise dose is zero. Restricted cubic splines with knots at the 10th, 50th and 90th percentiles of the explanatory variables were
used to investigate whether there was evidence for a nonlinear relationship. Fitting was performed by maximum likelihood, and nonlinearity was assessed using a likelihood ratio test. When the model was statistically significant at \( p<0.05 \), the minimum exercise dose to obtain a clinically meaningful benefit was estimated by the minimum value of the explanatory variable at which the estimated OR was less than 0.75. Linear models were presented unless the fit of the spline was significantly better. For outcomes or for subsets of studies where a meta-analysis was not possible, a narrative synthesis of the results was conducted, organised around each outcome. Unless otherwise specified, studies were not included in meta-analyses if data were reported incompletely (eg, SD, SE or number of cases/controls not provided), if data were adjusted for confounding factors or if the study did not include a non-exercising control group. In studies where data were included in the meta-analysis but additional information was available that could not be meta-analysed, the studies were incorporated in both the meta-analysis and narrative synthesis. Within each outcome, results were presented by study design.

**RESULTS**

**Study selection**

The initial search was not limited by language. However, the Guidelines Steering Committee decided to exclude studies published in languages other than English, Spanish or French for feasibility reasons. A PRISMA diagram of the search results, including reasons for exclusion, is shown in figure 1. A comprehensive list of excluded studies is presented in the online supplementary file.

**Study characteristics**

Overall, 52 unique studies (\( n=131,406 \) women) from 19 countries were included. There were 26 RCTs, 7 non-randomised intervention, 10 cohort, 6 cross-sectional and 3 case control
studies. Among the included exercise interventions, the frequency of exercise ranged from 1 to 7 days per week, the duration of exercise ranged from 20 to 75 min per session and the types of exercise included aerobic exercise, yoga, resistance training and pelvic floor muscle training. Additional details about the studies can be found in the online supplementary file. Possible diagnosis of depression or being high in anxiety was based on a cut-off score on a clinical scale. Severity of symptoms was the overall score on a clinical scale. State anxiety is an acute response to a perceived or real threat, while trait anxiety describes the inter-individual tendency to have high state anxiety in response to a threat.

Quality of evidence
Overall, the quality of evidence ranged from ‘very low’ to ‘high’ (see online supplementary tables 2 and 3). The most common reasons for downgrading the quality of evidence were (1) serious risk of bias that reduced the level of confidence in the observed effects and (2) indirectness of the interventions being assessed. Common sources of bias included poor or unreported compliance and inappropriate treatment of missing data when attrition rate was high. Publication bias was not observed among the analyses where it was possible to systematically assess publication bias using funnel plots.

Synthesis of data
The results of the meta-regression analysis are presented in the online supplementary file (Meta-regressions, online supplementary figures 31–34). Meta-regression analysis using linear and spline regression was conducted for each outcome. In each case, the spline model did not provide a significantly better fit than the linear model. Minimum exercise thresholds required to achieve a moderate effect size (SMD>0.4) were estimated. In order to achieve a moderate effect of exercise in reducing prenatal depressive symptoms, the following thresholds were identified: an exercise intensity of 4.0 METs/min (eg, light walking; online supplementary figure 31), exercise duration >49.9 min per session (online supplementary figure 32), exercise frequency >2.6 times per week (online supplementary figure 33) and volume of exercise per week >644 MET-min/week (eg, 150 min of moderate intensity exercise, such as brisk walking, water aerobics, stationary cycling, resistance training; online supplementary figure 34).

Prenatal depressive symptoms
Overall, there was ‘very low’ quality evidence from 21 RCTs (n=4267) regarding the association between prenatal exercise and depressive symptoms. The quality of evidence was downgraded from ‘high’ to ‘very low’ because of serious risk of bias, serious inconsistency and serious indirectness of the intervention. Overall, prenatal exercise was associated with a small reduction in severity of depressive symptoms compared with no exercise (pooled estimate based on 19 RCTs, n=3316; SMD: −0.23, 95% CI −0.36 to −0.09, I²=63%; figure 2). Two exercise-only interventions could not be included in the meta-analysis (see online supplementary file). In the first RCT, there was no influence of prenatal exercise on the severity of depressive symptoms during pregnancy. In contrast, in the second RCT, yoga (n=51) but not non-yoga (n=45) antenatal exercises was associated with an improvement in depressive symptoms.

Figure 2  Effects of prenatal exercise-only versus exercise plus co-interventions (randomised controlled trials) compared with control on prenatal depressive symptoms. Analyses conducted with a random effects model. IV, inverse variance.
Review

Sensitivity analysis
The pooled estimate for the exercise-only interventions was significantly different than the exercise+cointervention subgroups (p<0.0001). Specifically, exercise-only interventions had a small effect on reducing the severity of depressive symptoms compared with no exercise (13 RCTs, n=1076; SMD: −0.38, 95% CI −0.51 to −0.25, I²=10%; 'moderate' quality evidence, downgraded due to serious risk of bias; online supplementary figure 1). There was no statistically significant difference for those participating in exercise+cointerventions (online supplementary figure 1).

Subgroup analyses
There were no statistically significant subgroup differences identified (see figure 2, online supplementary figure 3–7).

Other study designs
Findings from four non-randomised interventions,38–41 one cohort,42 and three cross-sectional studies43–45 were in agreement with the findings from RCTs. In contrast, findings from three cohort studies46–48 and one case-control study49 were not significant (see online supplementary file for more details).

Prenatal depression
Overall, there was ‘very low’ quality evidence from eight RCTs (n=2481) regarding the association between prenatal exercise and prenatal depression. The quality of evidence was downgraded from ‘high’ to ‘very low’ because of serious risk of bias, serious inconsistency and serious indirectness of the intervention. Overall, prenatal exercise was associated with 45% lower odds of developing prenatal depression compared with no exercise (OR 0.55, 95% CI 0.34 to 0.90, I²=59%; figure 3).

Sensitivity analysis
The pooled estimate for the exercise-only interventions was significantly different than the exercise+cointervention subgroups (p<0.0001). Specifically, exercise-only interventions (all interventions included supervised exercise) reduced the odds of having depression by 67% (5 RCTs, n=683; OR: 0.33, 95% CI 0.21 to 0.53, I²=0%; ‘moderate’ quality evidence, downgraded due to serious risk of bias; figure 3). There was no statistically significant difference in odds of developing depression for those participating in exercise+cointerventions compared with a no-exercise control (figure 3).

Subgroup analyses
Stratification by exercise type showed no statistically significant difference in odds of developing prenatal depression (see online supplementary figure 13). The remaining a priori subgroup analyses could not be conducted as studies with the relevant subgroups did not exist.

Other study designs
Findings from one non-randomised intervention,39 one cohort study47 and one cross-sectional study50 were in agreement with the findings from RCTs (see online supplementary file for more details).

Postnatal depressive symptoms
Overall, there was ‘low’ quality evidence from seven RCTs (n=2795) regarding the association between prenatal exercise and severity of postnatal depressive symptoms. The quality of evidence was downgraded from ‘high’ to ‘low’ because of serious risk of bias and serious indirectness of the intervention. Overall, prenatal exercise was not associated with reduction in postnatal depressive symptoms compared with no exercise (SMD: 0.05, 95% CI −0.02 to 0.12, I²=0%; figure 4).

Other study designs
Findings from one non-randomised intervention,39 one cohort study47 and one cross-sectional study50 were in agreement with the findings from RCTs (see online supplementary file for more details).
Sensitivity analysis
The pooled estimate for the exercise-only interventions was not significantly different than the exercise+cointervention subgroups (p=0.26). Prenatal exercise-only interventions did not affect the severity of depressive symptoms in the postpartum period (figure 4).

Figure 4  Effects of prenatal exercise-only versus exercise+cointervention compared with control on postnatal depressive symptoms (randomised controlled trials). Analyses conducted with a random effects model. IV, inverse variance.

Subgroup analyses
Stratification by exercise type showed no statistically significant difference in odds of developing postnatal depressive symptoms. The remaining subgroup analyses on exercise-only intervention could not be conducted due to a lack of studies with relevant subgroups.

Other study designs
There was ‘very low’ quality evidence (downgraded due to serious risk of bias and serious inconsistency) from one cohort study (n=1305) showing a reduced odds of postnatal with prenatal exercise (OR 0.79, 95% CI 0.70 to 0.89; online supplementary figure 20).

Prenatal state anxiety symptoms
Overall, there was ‘very low’ quality evidence from eight RCTs (n=1785) regarding the association between prenatal exercise and prenatal state anxiety symptoms. The quality of evidence was downgraded from ‘high’ to ‘very low’ because of serious risk of bias, serious imprecision and serious indirectness of the intervention. Overall, prenatal exercise was not associated with a reduction in prenatal state anxiety symptoms compared with no exercise (pooled estimate based on seven RCTs, n=1689; SMD: 0.06, 95% CI −0.04 to 0.15, I²=0%; online supplementary figure 21). The one superiority trial that could not be included in the meta-analysis showed an improvement in state anxiety symptoms with yoga (n=51), but not with other forms of antenatal exercise (n=45).

Sensitivity analysis
The pooled estimate for the exercise-only interventions was not significantly different than the pooled estimate for the exercise+cointervention subgroups (p=0.12). Prenatal exercise-only interventions did not affect the severity of depressive symptoms in the postpartum period (online supplementary figure 18).
Review

Figure 5  Effects of prenatal exercise compared with control on prenatal state anxiety symptoms (randomised controlled trials). Subgroup analyses were conducted with studies including women with depression prior to the intervention with women who without depression (defined as ‘general population’). Analyses conducted with a random effects model. IV, inverse variance.

Subgroup analyses
The tests for subgroup differences performed for exercise-only interventions were not statistically significant (figure 5).

Other study designs
The findings from one non-randomised intervention and one cohort study were in agreement with the findings from RCT. In contrast, one cross-sectional study did not report a reduction in prenatal state anxiety symptoms with prenatal exercise. See online supplementary file for more details.

Prenatal state anxiety
Overall, there was ‘very low’ quality evidence from two RCTs of exercise+cointervention regarding the association between prenatal exercise and odds of prenatal state anxiety. The quality of evidence was downgraded from ‘high’ to ‘very low’ because of serious risk of bias, serious indirectness of the intervention and serious imprecision. Overall, prenatal exercise was not associated with lower odds of prenatal state anxiety compared with no exercise (OR: 1.12, 95%CI 0.85 to 1.48, I²=0%; online supplementary figure 22).

Sensitivity analysis
Sensitivity analysis could not be conducted as there were no exercise-only interventions.

Prenatal trait anxiety symptoms
Overall, there was ‘low’ quality evidence from two RCTs of exercise-only interventions regarding the association between prenatal exercise and severity of prenatal trait anxiety symptoms. The quality of evidence was downgraded from ‘high’ to ‘low’ because of serious risk of bias and serious imprecision. Overall, prenatal exercise was not associated with a decrease in prenatal trait anxiety symptoms compared with no exercise (SMD: −0.21, 95%CI −0.63 to 0.20, I²=0%; online supplementary figure 25). The one superiority trial that could not be included in the meta-analysis showed an improvement in trait anxiety symptoms with yoga (n=51), but not with other antenatal exercise (n=45).

Sensitivity analysis
Sensitivity analysis could not be conducted as there were no exercise+cointervention.

Other study designs
Findings from one non-randomised intervention and one case control study were in agreement with findings from RCTs (online supplementary figures 26 and 27), while a cross-sectional study demonstrated a reduction in the severity of prenatal trait anxiety symptoms in women who exercised during pregnancy compared with those who were inactive (n=203; SMD: −0.73, 95%CI −1.02 to −0.45; online supplementary figure 28).

Prenatal state anxiety
Overall, there was ‘very low’ quality evidence from one RCT of exercise+cointerventions regarding the association between prenatal exercise and prenatal trait anxiety. The quality of evidence was downgraded from ‘high’ to ‘very low’ because of serious risk of bias, serious indirectness and serious inconsistency. Overall, prenatal exercise was not associated with lower odds of prenatal trait anxiety compared with no exercise (OR: 0.75, 95%CI 0.36 to 1.56; online supplementary figure 29).

Postnatal state anxiety symptoms
Overall, there was ‘very low’ quality evidence from two RCTs regarding the association between prenatal exercise and postpartum state anxiety symptoms. The quality of evidence was downgraded from ‘high’ to ‘very low’ because of serious risk of bias, serious indirectness of the intervention and serious imprecision. Overall, prenatal exercise was not associated with reduction in postpartum state anxiety symptoms compared with no exercise (SMD: 0.01, 95%CI 0.10 to 0.12, I²=0%; online supplementary figure 30).

Sensitivity analysis
The pooled estimate for the exercise-only interventions was not significantly different than the pooled estimate for the exercise+cointervention subgroups (p=0.99).
exercise-only interventions did not affect the severity of postnatal state anxiety symptoms (online supplementary figure 30).

Postnatal state anxiety
Overall, there was ‘very low’ quality evidence from one exercise+cointervention RCT (n=1220)\(^{56}\) regarding the association between prenatal exercise and postnatal state anxiety. The quality of evidence was downgraded from ‘high’ to ‘very low’ because of serious risk of bias, serious indirectness and serious inconsistency. Overall, prenatal exercise was not associated with lower odds of postnatal state anxiety compared with no exercise (OR: 1.28, 95%CI 0.91 to 1.80; online supplementary figure 31).

DISCUSSION
The main finding of our systematic review was that exercise-only interventions were associated with a reduction in the severity of prenatal depressive symptoms (SMD −0.39), and this reduction was greatest in women who were supervised during exercise. Pregnant women need to accumulate at least 644 MET-min/week of moderate-to-vigorous intensity exercise (eg, 150 min of moderate intensity exercise, such as brisk walking, water aerobics, stationary cycling, resistance training) in order to achieve a moderate effect size in the reduction in the severity of prenatal depressive symptoms. There appears to be a dose-dependent association of this benefit. A greater volume of exercise was associated with a greater reduction in the severity of prenatal depressive symptoms. Further, exercise-only interventions were associated with a 67% reduction in the odds of developing prenatal depression.

Postnatal depression—symptoms and diagnosis—was not reduced by prenatal exercise. There was no association between prenatal exercise and anxiety or the severity of anxiety symptoms in the prenatal or postnatal period.

SENSITIVITY ANALYSIS
Sensitivity analyses comparing trials with cointerventions (diet+exercise; education classes+exercise; smoking cessation programme+exercise) against trials without cointerventions lowered heterogeneity within each group (11% and 0%, respectively). There was a significant difference between the two types of interventions; severity of depressive symptoms was lower postintervention in standalone exercise trials than for trials including cointerventions. The lack of improvement in depressive symptoms in response to a co-intervention may have been due to the fact that women were required to change multiple behaviours at once, in addition to responding to the demands of pregnancy.\(^{64}\) Thus, setting realistic and incremental goals may be important considerations when prescribing exercise in the treatment of depression.\(^{63}\)

Supervised exercise was associated with a moderate reduction in the severity of depressive symptoms. This may be partially explained by the social support women in supervised interventions received; low social support is a known correlate of depression.\(^{16}\) It has been suggested that supervision enhances compliance and effectiveness of interventions.\(^{64}\) However, because adherence was either not reported by the authors in a standard manner or not reported at all, we were not able to analyse the effects of the exercise-only interventions in compliant versus non-compliant women to focus on the issue of compliance.

What mechanisms may underpin these epidemiological findings?
Rauff and Downs suggested that improvement in body image may be a factor mediating the positive relationship between exercise and reduced depression or depressive symptoms during pregnancy.\(^{64}\) Physiological mechanisms have also been proposed. In non-pregnant populations, aerobic exercise may normalise dopamine-serotonin levels and reduce stress, resulting in a subsequent decrease in anxiety and depression.\(^{65} 66\) Depression has been suggested to result, at least in part, from dysregulation of biochemical and neurophysiological function including attenuated release of norepinephrine, serotonin and dopamine.\(^{67}\) Exercise may counteract these effects by increasing the release of dopamine and serotonin as well as blunting the release of cortisol in response to stress.\(^{68}\)

In our systematic review of exercise-only interventions, both yoga and aerobic exercise had a small effect size in reducing the severity of depressive symptoms. However, these data were from a small number of trials and women; thus, additional studies examining the potential impact of different types of exercise are warranted.

Is there a carryover benefit into the postpartum period?
Prenatal exercise-only interventions had no impact on depressive symptoms or the incidence of depression during the postpartum period. However, a recent meta-analysis suggested that exercise interventions initiated during the postpartum period (at least 4 weeks after delivery) reduced the severity of depressive symptoms.\(^{20}\)

We highlight that no trials examined populations with diagnosed clinical anxiety disorders. Our meta-analysis does not extend to the efficacy of exercise in the treatment of anxiety disorders (generalised anxiety disorder, panic disorder, social anxiety disorder, among others).

Strengths of this review include our broad inclusion criteria covering all studies types (except case studies), grey literature and three languages (English, French and Spanish from 19 countries). We applied rigorous methodology (GRADE, PRISMA) to the systematic review process and evaluation of the quality of evidence. Subgroup analyses allowed identification of possible sources of heterogeneity; however, statistical heterogeneity was still moderate within subgroups. There was also considerable variability regarding the tools used to measure depression/depressive symptoms, with the majority of trials using either the CES-D or EPDS. As noted by Gong (2015), the CES-D, while a validated measure of perinatal depression, may misattribute some of the somatic symptoms associated with pregnancy (lack of energy, tiredness) to depression, which may have underestimated the effectiveness of exercise interventions. Five trials examined prenatal exercise and depression or anxiety and showed considerable baseline differences in depression and anxiety. However, an exploratory sensitivity analysis performed without these trials had no significant impact on findings. Finally, some of the trials included in our review included women who were using antidepressants, an important confounding variable, that may have influenced the findings.

CONCLUSION
Overall, prenatal exercise reduced the odds of prenatal depression and the severity of depressive symptoms. This positive effect of prenatal exercise did not extend to the postpartum period or prenatal or postnatal anxiety.
What are the new findings?

- Exercise-only interventions reduced the severity of prenatal depressive symptoms with an effect size similar to that found for psychological treatments in depressed prenatal women.
- Odds of prenatal depression were reduced by 67% in women who engaged in prenatal exercise.
- The benefits of prenatal exercise on depression or depressive symptoms did not extend into the postpartum period.
- To achieve at least a moderate reduction in the severity of prenatal depressive symptoms, pregnant women needed to accumulate at least 644 MET-min/week of exercise (eg, 150 min of moderate intensity exercise, such as brisk walking, water aerobics, stationary cycling, resistance training).

What is already known?

- Depression and anxiety during and following pregnancy impact the short-term and long-term health issues for mother and child.
- Exercise prevents and treats anxiety and depression in non-pregnant populations. Whether exercise prevents and treats mental health issues during pregnancy is poorly understood.

REFERENCES

1. Friedl MI. Depression is the leading cause of disability around the world. JAMA 2017;317:1517.