Exercise for the prevention and treatment of low back, pelvic girdle and lumbopelvic pain during pregnancy: a systematic review and meta-analysis

Margie H Davenport,1 Andree-Anne Marchand,2 Michelle F Mottola,3 Veronica J Poitras,4 Casey E Gray,5 Alejandra Jaramillo Garcia,4 Nick Barrowman,6 Frances Sobierajski,1 Marina James,1 Victoria L Meah,7 Rachel J Skow,1 Laurel Riske,1 Megan Nuspl,8 Tanya S Nagpal,3 Anne Courbalay,2 Linda G Slater,9 Kristi B Adamo,10 Gregory A Davies,11 Ruben Barakat,12 Stephanie-May Ruchat13

ABSTRACT

Objective The purpose of this review was to investigate the relationship between prenatal exercise, and low back (LB), pelvic girdle (PGP) and lumbopelvic (LBPP) pain.

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 eligible (except case studies and reviews) 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, alone ["exercise-only"] or in combination with other intervention components [eg, dietary; "exercise + co-intervention"]), comparator (no exercise or different frequency, intensity, duration, volume and type of exercise) and outcome (prevalence and symptom severity of LBP, PGP and LBPP).

Results The analyses included data from 32 studies (n=52,297 pregnant women). 'Very low' to 'moderate' quality evidence from 13 randomised controlled trials (RCTs) showed prenatal exercise did not reduce the odds of suffering from LB, PGP and LBPP either in pregnancy or the postpartum period. However, 'very low' to 'moderate' quality evidence from 15 RCTs identified lower pain severity during pregnancy and the early postpartum period in women who exercised during pregnancy (standardised mean difference −1.03, 95% CI −1.58, −0.48) compared with those who did not exercise. These findings were supported by 'very low' quality evidence from other study designs.

Conclusion Compared with not exercising, prenatal exercise decreased the severity of LBP, PGP or LBPP during and following pregnancy but did not decrease the odds of any of these conditions at any time point.

INTRODUCTION

Approximately 50% of women experience low back (LB) or pelvic girdle (PGP) pain during pregnancy; 25% continue to experience pain 1 year after delivery. A 10-year follow-up study reported that 1 in 10 women with PGP in pregnancy have severe consequences up to 11 years postpartum.1-3 LBP is pain or discomfort located between the 12th rib and the gluteal fold, and PGP has been defined as 'pain experienced between the posterior iliac crest and the gluteal fold, particularly in the vicinity of the sacroiliac joints'.4 Despite the fact that both conditions are considered distinct entities, the concomitance of LBP and PGP, herein referred to as lumbopelvic pain (LBPP), puts a greater burden on pregnant women regarding health quality and daily functioning.5 With repercussions such as disruption of sleep, social and sexual life, work capacity and increased psychological stress,6 it is not surprising that pregnant women experiencing PGP have also been reported to be less likely to exercise regularly during pregnancy.7

In the general population, 'moderate' quality evidence suggests exercise has a small positive effect on the severity of LBP compared with usual care, which is comparable with the effectiveness of other non-pharmacological approaches recommended for the management of acute or chronic LBP.8 However, compared with other cost effective non-pharmacological treatments, such as interdisciplinary rehabilitation, acupuncture, spinal manipulation or cognitive behavioural therapy, exercise is easily accessible as part of a self-management strategy, can require minimal equipment and can be...
performed at home. Previous national and international guidelines for exercise during pregnancy endorsed the benefits of exercise during pregnancy in terms of fitness, overall well being and decreased risk of developing pregnancy related complications. However, in these guidelines, exercise was not considered within the context of pregnancy related pain. The European guidelines for the diagnosis and treatment of PGP, published in 2008, were the first to mention that exercise should be recommended for pregnant women, with a focus on activities of daily living and avoidance of maladaptive movement patterns. The Cochrane Review published in 2015 by Liddle et al supports this recommendation, concluding that ‘exercise may reduce pregnancy related low back pain’. However, research gaps remain because their findings were based on a limited number of randomised controlled trials (RCTs) containing ‘low’ quality evidence, the effect of prenatal exercise on PGP was not conclusive and the effects of prenatal exercise on LBP, PGP and LBPP during the postpartum period were not examined.

The present systematic review and meta-analysis was conducted as 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 review was to examine the effect of prenatal exercise (in terms of frequency, intensity, type and volume) on maternal LBP, PGP and LBPP during pregnancy and in the postpartum period.

METHODS
In October 2015, the Guidelines Consensus Panel assembled to identify priority outcomes for the development of the Guidelines. The panel included researchers, methodological experts, a fitness professional and representatives from the Canadian Society for Exercise Physiology, the Society of Obstetricians and Gynaecologists of Canada, the College of Family Physicians of Canada, the Canadian Association of Midwives, the Canadian Academy of Sport and Exercise Medicine, Exercise is Medicine Canada and a representative health unit (the Middlesex-London Health Unit). The Guidelines Consensus Panel selected 20 ‘critical’ and 17 ‘important’ outcomes related to prenatal exercise and maternal or fetal health. Prevalence and severity of LBP, PGP and LBPP during pregnancy and in the postpartum period were deemed ‘important’ outcomes. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement and checklist were used to guide this systematic review and meta-analysis.

Protocol and registration
Two systematic reviews were undertaken to investigate the impact of prenatal exercise on fetal and maternal health outcomes and records identified through both processes were considered for inclusion in the current review. Each review was registered a priori with the International Prospective Register of Systematic Reviews (PROSPERO; fetal health: CRD42016029869; maternal health: CRD42016032376). Although the relationships between prenatal exercise and the prevalence or severity of maternal LBP, PGP or LBPP are primarily examined in studies related to maternal health, records retrieved from both of these searches were evaluated for inclusion in the current review.

Eligibility criteria
This study followed the participants, interventions, comparisons, outcomes, and study design (PICOS) framework.

Population
The population of interest was pregnant women without absolute or relative contraindications to exercise (according to the guidelines of the Canadian Society for Exercise Physiology and the American Congress of Obstetricians and Gynaecologists). Absolute contraindications to exercise were defined as: ruptured membranes, premature labour, persistent second or third trimester bleeding, placenta previa, pre-eclampsia, 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.

Intervention (exposure)
The intervention/exposure was objectively or subjectively measured prenatal exercise of any frequency, intensity, duration, volume or type. Prenatal exercise could be acute (ie, a single exercise session) or habitual (ie, usual activity). Interventions that consisted of exercise alone (termed ‘exercise only’ interventions) or exercise combined with other interventions (eg, diet; termed ‘exercise+co-interventions’) were considered. Studies were not eligible if exercise began after the initiation of labour. Although exercise is a subtype of physical activity, the terms are used interchangeably in this review. Exercise and physical activity were defined as any bodily movement generated by skeletal muscles that resulted in energy expenditure above resting levels.

Comparison
Comparators that were eligible were: no exercise; different frequency, intensity, duration, volume or type of exercise; different duration of intervention; or exercise in a different trimester.

Outcome
Relevant outcomes were prevalence and severity of maternal LBP, PGP or LBPP (as defined by the authors) during pregnancy and the postpartum period (up to 1 year following delivery).

Study design
Studies of any design were eligible, with the exception of case studies (n=1), narrative syntheses and systematic reviews.

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 and Adolescent Studies, ERIC, Sport Discus, ClinicalTrials.gov and the Trip Database, up to 6 January 2017 (see the online supplement for the complete search strategies).

Study selection and data extraction
The titles and abstracts of all retrieved articles were independently screened by two reviewers. Abstracts that were deemed to have met the inclusion criteria by at least one reviewer were automatically retrieved as full text articles. Full text articles were independently screened by two reviewers for the relevant population,
intervention, comparators and outcomes before data extraction. For studies where at least one reviewer recommended exclusion, further review was conducted by MHD and/or S-MR for a final decision on exclusion. In the event of a disagreement that could not be resolved through discussion, the characteristics of the study were presented to the Guidelines Steering Committee who conducted the systematic reviews (MHD, MFM, S-MR, CEG, VJE, AJG and NB) and a final decision regarding inclusion/exclusion was made by consensus. Studies identified by the maternal and fetal search strategies were imported into DistillerSR for de-duplication and data extraction, and were subsequently considered as one review.

Data extraction tables were created in DistillerSR in consultation with methodological experts and the Guidelines Steering Committee. Data from records that met the inclusion criteria were extracted by one person and independently verified by a content expert (MHD, MFM or S-MR). For studies where multiple publications existed, the most recent or complete publication was selected as the ‘parent’ paper; however, relevant data from all publications were extracted. Extracted data were study characteristics (ie, author, year, study design and country), characteristics of the population (eg, number of participants, age, pre-pregnancy body mass index (BMI), parity and pregnancy complications including pre-eclampsia, gestational hypertension and gestational diabetes), intervention/exposure (prescribed and/or measured exercise frequency, intensity, time, type and volume, duration of the intervention and measurement tool) and outcomes (prevalence and/or severity of LBP, PGP or LBPP). In cases where data were unavailable for extraction, the authors were contacted for additional information (see the online supplement table 1, for included study characteristics).

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.22 Evidence from RCTs was considered ‘high’ quality and was graded down if there was a concern with risk of bias,23 indirectness,24 inconsistency,25 imprecision26 or risk of publication bias27 because these factors reduce the level of confidence in the observed effects. Evidence from all non-randomised interventions and observational studies was considered ‘low’ quality and, if there was no cause to downgrading, was upgraded if applicable according to the GRADE criteria (eg, large magnitude of effect, evidence of dose response).28 Specifically, the risk of bias in RCTs and intervention studies was assessed following the Cochrane Handbook,29 and the risk of bias in observational studies was assessed using the characteristics recommended by Guyatt et al.24 consistent with systematic reviews conducted to support previous health behaviour guidelines.30 31 All studies (RCTs, intervention studies and observational studies) were screened for potential sources of bias, including selection bias (RCT/intervention: inadequate randomisation procedure; observational: inappropriate sampling), reporting bias (selective/incomplete outcome reporting), performance bias (RCT/intervention: compliance with the intervention; observational: flawed measurement of exposure), detection bias (flawed measurement of outcome), attrition bias (incomplete follow-up, high loss to follow-up) and ‘other’ sources of bias. The risk of bias across studies was considered ‘serious’ when studies contributing the most to the pooled result (assessed using weight (%)) given in forest plots or sample size in studies that were narratively synthesised) presented ‘high’ risk of bias. Specifically, the studies with the greatest contribution to the pooled result were determined as follows: (1) those that had the greatest individual per cent weight in the meta-analyses, when taken together, contributed to >50% of the weight of the pooled estimate and (2) the 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.

Due to the nature of physical activity interventions, it is not possible to blind participants to group allocation, and selection risk of bias was rated as ‘low’ if this was the only source of bias identified. Performance bias was rated as ‘high’ when <60% of participants performed 100% of the prescribed exercise sessions or attended 100% of the counselling sessions (defined as low compliance) or when compliance with the intervention was not reported. Attrition bias was rated as ‘high’ risk when >10% of participants dropped out of the study for any reason, and intention to treat analysis was not used.

Inconsistency across studies was considered serious when heterogeneity was high (I^2 ≥50%) or when only one study was assessed (I^2 unavailable). Indirectness was considered serious when interventions included both exercise and additional components (ie, exercise+co-interventions, or exercise only and exercise+co-interventions combined in analyses). Imprecision was considered serious when the 95% CI crossed the line of no effect, and was wide, such that the 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 assessed, imprecision was not considered serious because inconsistency was already considered serious for this reason. Finally, publication bias was assessed if possible (ie, at least 10 studies were included in the forest plot) via funnel plots (see online supplement, figures 12 and 13). If there were fewer than 10 studies, publication bias was deemed non-estimable and not rated down. Original plans for two people to independently assess the quality of the evidence across each health outcome were amended for feasibility reasons. As such, one reviewer evaluated the quality of the evidence and a second person checked the GRADE tables as a quality control measure. GRADE tables are presented in the online supplement, tables 2 and 3.

Statistical analysis and narrative synthesis
Statistical analyses were conducted using Review Manager V.5.3. (Cochrane Collaboration, Copenhagen, Denmark). Statistical significance was set at P<0.05. For all dichotomous outcomes, odds ratio (OR) were calculated. Inverse variance weighting was applied to obtain OR using a random effects model. For continuous outcomes, mean differences (MD) between the intervention and control groups were examined, and weighted mean differences were calculated using a random effects model. As severity of pain was assessed using multiple tools, standardised mean differences (SMD) were calculated when different tools were used for a single outcome. SMD effect sizes were calculated in Review Manager V.5.3 using Hedges’ g method (similar to Cohen’s d). Effect sizes of 0.2, 0.4 and 0.8 were considered small, moderate and large, respectively.32 Meta-analyses were performed separately by study design. A staged approach was utilised to determine if there was sufficient evidence from high quality study designs (ie, RCTs) to inform the Guidelines, or if lower quality study designs needed to be examined. If meta-analyses of RCTs contained data from fewer than 2000 women, the impact of prenatal exercise on the specific outcome was examined further utilising observational evidence (first non-randomised interventions; if 2000
women were not reached after adding non-randomised interventions, we used cohort, cross sectional and case control studies). For RCTs and non-randomised interventions, sensitivity analyses were performed to evaluate whether the effect for all types of pain (LBP, PGP and LBPP combined, whenever available) was different when examining evidence from exercise only interventions (including standard care) versus exercise+co-interventions. A priori determined subgroup analyses were conducted when possible for exercise only interventions and observational studies. These subgroups were: (1) women diagnosed with diabetes (gestational, type 1 or type 2) compared with women without diabetes (named ‘general population’); (2) women with pre-pregnancy overweight or obesity status (mean BMI >25.0 kg/m²) compared with women who were of various BMI (mean BMI <25 kg/m², which may include some individuals with BMI >25.0 kg/m²; named ‘general population’); (3) women >35 years of age compared with women <35 years of age; or (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 for inclusion into the a priori subgroups, then a third group called ‘unspecified’ was created. A priori subgroup analyses were also conducted for exercise only RCTs to identify whether relationships between exercise and outcomes differed depending on the type of exercise (eg, aerobic exercise, resistance training or yoga). Due to feasibility, these subgroup analyses were only conducted for outcomes rated as ‘critical’.

The per cent of total variability that was attributable to between study heterogeneity (ie, not to chance) was expressed using the I² statistic. In the case of I² >50%, heterogeneity was explored further with additional subgroup analyses, and the overall result was presented using the random effects model. In the particular case of this review additional post hoc analyses were conducted to assess whether exercise only interventions had different effects according to types of pain assessed individually (whenever data were available). For the studies that considered multiple pain presentations but did not break down the results by pain type, results were categorised as ‘mixed pain presentations’. In studies where there were 0% or 100% of events in the intervention or control group, 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.29

In order to identify a clinically meaningful decrease in pain, dose–response meta-regression33–35 was carried out by weighted no-intercept regression of log OR with a random effects for study, using the metafor package in R (V.3.4.1).36 It was determined that an accepted cut off point for a clinically meaningful decrease does not exist in the literature. As such, a reduction of 25% was chosen based on expert opinion. 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 variable37 were used to investigate whether there was evidence for a non-linear relationship. Fitting was performed by maximum likelihood, and non-linearity 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 significant benefit was estimated by the minimum value of the explanatory variable at which the estimated OR was <0.75. If there were fewer than 10 studies of similar design, meta-regression analysis was not conducted.

For outcomes where a meta-analysis was not possible, the results were presented as a narrative synthesis, structured around each outcome. Within each outcome, results were organised by study design. 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 the data were adjusted for confounding factors or if the study did not include a non-exercise 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 included in both the meta-analysis and narrative synthesis.

RESULTS
Study selection
Although the initial search was not limited by language, 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 supplement. Consistent with the planned staged approach, when fewer than 2000 participants were represented in included RCTs, data were considered from other study designs.

Study characteristics
Overall, 32 unique studies (n=52,297 women) from 14 countries over four continents (Asia, Europe, Americas and Africa) were included. There were 23 RCTs (including 13 exercise only interventions and 10 exercise+co-interventions), 5 non-randomised interventions, 3 cohort studies and 1 case control study. The co-interventions consisted of education on pregnancy related topics, including the role of pelvic floor muscles in pregnancy and postpartum, anatomical and physiological changes occurring during pregnancy, nutrition, non-pharmacological pain prevention and management strategies, labour and ergonomics. The exercise components of the interventions included yoga, aerobic exercise, general muscle strengthening or muscle strengthening specific to one body region and the combination of aerobic and resistance training. The majority of the interventions started during the second trimester of pregnancy with most of the studies ending in the late third trimester. The frequency of exercise ranged from 1 to 14 times per week, the duration of exercise ranged from 20 to 75 min per session and the intensity of exercise ranged from low to vigorous (see the online supplement for more details).

Synthesis of the data
The a priori planned subgroup analyses were not conducted for this meta-analysis because the included studies did not provide sufficient detail to allow them, and because this meta-analysis does not include outcomes rated as ‘critical’. Dose–response meta-regressions were not carried out because we did not have 10 studies of similar design to include in the analysis.

Odds of LBP, PGP and LBPP during pregnancy
There was ‘very low’ quality evidence from 13 RCTs (n=2253) regarding the association between prenatal exercise and odds of any type of pain (LBP, PGP or LBPP) during pregnancy.38–40 The quality of evidence was downgraded from ‘high’ to ‘very low’ due to serious risk of bias, serious indirectness of the interventions and serious imprecision. Overall, prenatal exercise was not associated with a lower odds of pain during pregnancy compared with no exercise (pooled estimate based on 12 RCTs, n=1987; OR 0.78, 95% CI 0.6, 1.02, I²=22%) (figure 2).38–40 42–50 One additional study by Granath et al41 that could not be included in the meta-analysis (no numerical data provided) indicated that women
participating in water-based exercise (n=132) were less likely to report LBP during pregnancy (P=0.04) than those participating in land-based exercise (n=134), whereas the rate of PGP did not differ between the two groups (see online supplement table 1).

Sensitivity analysis
The pooled estimate for the exercise only interventions was not significantly different than the pooled estimate for the exercise+co-interventions (P=0.24). There was ‘very low’ quality evidence (downgraded due to very serious risk of bias, serious inconsistency and serious imprecision) showing that exercise only interventions did not reduce the odds of any type of pain located in the lumbopelvic region during pregnancy (pooled estimate based on six studies, n=343; OR 0.51, 95% CI 0.22, 1.16. I²=52%) (figure 2). Similar results were found for exercise+co-interventions (figure 2).

Post hoc analysis
The test for differences between different types of pain performed for exercise only interventions was not statistically significant (P=0.68; see online supplement, figure 3).

Severity of symptoms for LBP, PGP and LBPP during pregnancy
There was ‘very low’ quality evidence from 14 RCTs (n=1188) indicating an inverse association between prenatal exercise and severity of any type of pain (LBP, PGP or LBPP) during pregnancy. The pooled estimate was based on 10 RCTs (n=784; SMD −1.03 (large effect size), 95% CI −1.58, −0.48, I²=92%) (figure 3). The quality of evidence was downgraded from ‘high’ to ‘very low’ because of very serious risk of bias, serious inconsistency and serious indirectness of the interventions. Four additional RCTs that could not be included in the meta-analysis due to lack of a non-exercise control group (superiority trials) or incomplete reporting of data also indicated that prenatal exercise was associated with lower severity of any type of pain (LBP, PGP or LBPP) during pregnancy (see online supplement table 1). A superiority trial by Peterson et al reported that short periods of strengthening and pelvic tilt exercises (n=22) performed five times per week decreased LBP severity (3.9 (1.5) vs 2.4 (1.8) on an 11 point numeric pain rating scale) similarly to a neuro-emotional technique (n=20) and spinal manipulative therapy (n=15). Although clinically significant, the decrease was

Figure 1 Flow diagram of study selection.
not statistically significant. Kilhstrand et al reported that 30 min of water aerobics (intervention group, n=129) performed once a week in the second half of pregnancy reduced average weekly LBP intensity during the 31st week of gestation, as well as between the 33rd and 38th weeks of gestation (no statistics provided) compared with no exercise (control group, n=128). Kluge et al reported that combining supervised and home based graded exercise sessions (n=26) for 10 weeks decreased LBPP severity compared with no exercise (n=24) (change score for the intervention group: -11.5/60, P<0.01, change score for the control group: -2.0/60, P=0.89). Finally, a superiority trial by Gupta et al reported greater improvement in LBPP symptoms in women who performed stabilisation exercise and physiotherapy (n=20) compared with those who did standard physiotherapy alone (n=20) (mean scores difference -2.25/10, P=0.037).

Sensitivity analysis
The pooled estimate for the exercise only interventions was not significantly different than the pooled estimate for the exercise + co-interventions (P=0.05). Specifically, there was ‘very low’ quality evidence (downgraded due to very serious risk of bias and serious inconsistency) from eight RCTs showing that exercise only interventions reduced the severity of LBP during pregnancy (pooled estimate based on seven RCTs, n=552; SMD -1.43 (large effect size), 95% CI -2.29, -0.58, I²=94%) (figure 3). In contrast, there was ‘very low’ quality evidence (downgraded due to serious risk of bias, serious inconsistency, serious indirectness and serious imprecision) from six RCTs reporting that exercise + co-interventions did not reduce the severity of any type of pain (LBP, PGP or LBPP) during pregnancy. The pooled estimate was based on three RCTs (n=232;
SMD $-0.45$ (medium effect size), $95\% CI -0.94, 0.03, I^2=68\%$) (figure 3).\textsuperscript{44,47,58}

Post hoc analysis

Post hoc analyses on exercise only interventions could not be conducted since exercise only interventions reported only on LBP.

Odds of LBP, PGP and LBPP during the postpartum period

There was ‘low’ quality evidence from three RCTs ($n=491$) indicating no association between prenatal exercise and odds of any type of pain (LBP, PGP or LBPP) during the postpartum period (OR $0.89, 95\% CI 0.51, 1.56, I^2=27\%$) (figure 4).\textsuperscript{42,43,46} The quality of the evidence was downgraded from ‘high’ to ‘low’ because of serious indirectness of the intervention and serious imprecision.

Sensitivity analysis

The pooled estimate for the exercise only interventions was not significantly different than the pooled estimate for the exercise+co-interventions ($P=0.70$). There was ‘low’ quality evidence (downgraded due to serious indirectness and serious imprecision) to ‘moderate’ quality evidence (downgraded due to serious inconsistency) that neither exercise only interventions nor exercise+co-interventions affected the odds of any type of postpartum pain (LBP, PGP or LBPP) (one exercise only intervention, $n=92$; OR $0.75, 95\% CI 0.19, 2.94$\textsuperscript{42}; two exercise+co-interventions, $n=399$; OR $1.01, 95\% CI 0.47, 2.18, I^2=49\%$) (figure 4).\textsuperscript{45,46}

Post hoc analysis

The test for subgroup differences performed for exercise only interventions was not statistically significant ($P=0.21$; see online supplement, figure 8).

Severity of symptoms of LBP, PGP and LBPP during the postpartum period

There was ‘low’ quality evidence from one RCT that could only be reported narratively (incomplete reporting of data) showing that 30 min of water aerobics performed once a week ($n=129$) in the second half of pregnancy reduced LBP intensity in the first week postpartum ($P=0.034$) compared with no exercise ($n=128$).\textsuperscript{36} The quality of the evidence was downgraded from ‘high’ to ‘low’ because of serious risk of bias and serious inconsistency. There were no studies looking at the effect of prenatal exercise on other types of pain during the postpartum period.

According to the staged approach, observational study designs were examined. The results for these studies are presented in the online supplement. In summary, results from one non-randomised intervention\textsuperscript{61} and one cohort study\textsuperscript{62} were in agreement with those of RCTs and indicated that prenatal exercise performed in a variety of formats did not reduce the odds of LBP, PGP or LBPP either during pregnancy or in the early postpartum period.

Similar to non-specific LBP in the general population, risk factors for pregnancy related LBP and PGP are believed to be multifactorial (eg, increased joint laxity, displacement of gravity centre and increased axial loading, and vascular changes) and the exact pathogenesis remains unclear.\textsuperscript{67} The complex nature and unpredictable course of the condition further complicate the prevention of first time symptom onset. Considering this
The causes of LBP and PGP are multifactorial, including biomechanical factors such as muscle imbalances or altered alignment, and psychosocial factors such as stress or anxiety. 64,65 Exercise has been evaluated as a means to decrease pain severity during pregnancy; however, the results are mixed. 66 The current meta-analysis provided some evidence that different types of exercise performed alone or in combination, such as yoga, general or specific strengthening exercise and aerobic exercise performed anywhere from once per week to once per day, can reduce pain severity in non-specific LBP pain intensity is estimated to be at least 15–20% (specific estimates vary based on length of symptoms, baseline scores and tools used). 69 Considering that LBPP affects more than half of pregnant women and is a debilitating condition that is associated with disability, depression, reduced quality of life and higher prevalence of sick leave during pregnancy, 70 exercise may offer a cost effective, self-management strategy option to expecting mothers as part of a multimodal approach to decrease symptom severity.

Although the mechanisms through which exercise may reduce pain severity remain unclear, it is suggested that being physically active lessens the degree of biomechanical change occurring as pregnancy advances, such as decreasing the load on the spine, increasing joint stabilisation and contributing to better spinal alignment and segmental motion. From a more general standpoint, exercise may help reverse trunk muscle imbalance 70 or initiate a pain desensitisation process leading to increased pain detection threshold. 71

Only four studies investigated the effect of prenatal exercise on pain prevalence beyond pregnancy, and only one study looked at pain severity in the first week postpartum. Despite the limited evidence, the results are in favour of prenatal exercise as a means to decrease pain severity during the postpartum period. However, given the potential of LBPP to become debilitating when persisting beyond delivery, 72 and considering that only one study has evaluated the impact of prenatal exercise on pain severity at 1 week postpartum, further studies are needed to examine the effect of prenatal exercise on pain severity during this time period.

It should be noted that the definitions used by authors to classify LBPP, PGP or LBPP were often undisclosed. Similarly, the methods used to ascertain the presence of symptoms ranged from self-reported to objective testing (such as pain provocative tests; see online supplement for more details) performed during clinical evaluation, which likely resulted in high heterogeneity in the populations, or even possibly misclassified and improperly included women. Also, a small number of studies provided results of pain severity using outcomes designed to assess LBPP with disability rather than pain, which lowers the confidence in the overall estimates. In addition, the high heterogeneity detected in the analyses examining the effect of prenatal exercise on pain severity in combination with sample sizes, high dropout rates, lack of compliance with the exercise interventions and poor reporting on the use of co-interventions. Studies not included in the meta-analysis, that were narratively summarised included those reporting median scores and ranges rather than means and SD, those omitting raw data scores or those lacking a control group. 41,54-65 Future studies would benefit from designing and presenting studies in accordance with accepted reporting standards (eg, CONSORT) to strengthen the quality of the evidence.

This systematic review has several strengths. Rigorous methodological standards (GRADE) were used to guide the systematic review process. Grey literature was examined, allowing for identification of information on outcomes that would have otherwise been absent; articles published in different languages and studies of all designs were included. When it was not possible to include results in a meta-analysis, results were reported narratively. The broad inclusion criteria makes the present systematic review the most comprehensive to date. Fourteen countries from four continents were represented in the included studies. A limitation of our systematic review was that data extraction was not done in duplicate; rather, data were extracted by one individual and subsequently verified by a content expert. We were also unable to identify evidence based cut off points for clinically meaningful changes in study outcomes. Accordingly, it is possible that the results may have over- or underestimated the relevance of the findings.

CONCLUSION
Exercise initiated during pregnancy was not effective in decreasing the prenatal or postnatal prevalence of LBP, PGP or LBPP. In contrast, prenatal exercise had a large effect on decreasing the severity of LBP, PGP and LBPP during pregnancy; one study supported this finding for the postpartum period, although others showed no association. Further research is needed in order to identify the best exercise modalities, as well as the best period during pregnancy to initiate an exercise intervention, in order to optimise the treatment of maternal LBP, PGP and LBPP.
Acknowledgements The authors wish to acknowledge Mary Duggan, Canadian Society for Exercise Physiology, who is the primary knowledge user for the Canadian Institute of Health Research Knowledge Synthesis Grant. We would like to thank Meghan Sebastianski (Alberta SPOR SUPPORT Unit Knowledge Translation Platform), University of Alberta, for her assistance with the meta-analyses.

Contributors MHD, S-MR, MMF, GAD and KBA contributed to the conception of the study. MHD, S-MR, MMF, GAD, KBA, AIJ, NB, VJP, CEG, LGS and RB contributed to the design of the study and development of the search strategy. LGS conducted the systematic search. FS, MJ, VLM, RS, LR, MN, TSN, A-AM and AC completed the acquisition of the data. MHD, NB, S-MR and A-AM performed the data analysis. All authors assisted with the interpretation. A-AM, S-MR and MHD were the principal writers of the manuscript. All authors contributed to the drafting and revision of the final article. All authors approved the final submitted version of the manuscript.

Funding This project was funded by a Canadian Institute of Health Research Knowledge Synthesis Grant. MHD is funded by an Advancing Women’s Heart Health Initiative New Investigator Award supported by Health Canada and the Heart and Stroke Foundation of Canada. A-AM is funded by a Fonds de Recherche du Québec–Santé Doctoral Research Award. RS is funded by a Canadian Institutes for Health Research Doctoral Research Award.

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES
Review


