Psychological Medicine

cambridge.org/psm

Physical activity and prenatal depression: going beyond statistical significance by assessing the impact of reliable and clinical significant change

Marina Vargas-Terrones[1], Taniya S. Nagpal[2,3], Maria Perales[4,5], Harry Prapavessis[3], Michelle F. Mottola[2,6,7] and Ruben Barakat[1]

[1]AFIPE Research Group, Faculty of Sciences for Physical Activity and Sport, INEF, Universidad Politécnica de Madrid (UPM), Madrid, Spain; [2]R. Samuel McLaughlin Foundation-Exercise and Pregnancy Laboratory, School of Kinesiology, The University of Western Ontario, London, Ontario, Canada; [3]Exercise and Health Psychology Laboratory, School of Kinesiology, The University of Western Ontario, London, Ontario, Canada; [4]Camilo José Cela University, Madrid, Spain; [5]Research Institute of Hospital 12 de Octubre (i + 12), Madrid, Spain; [6]Department of Anatomy & Cell Biology, Schulich School of Medicine & Dentistry, The University of Western Ontario, London, Ontario, Canada and [7]Children's Health Research Institute, The University of Western Ontario, London, Ontario, Canada

Original Article

Cite this article: Vargas-Terrones M, Nagpal TS, Perales M, Prapavessis H, Mottola MF, Barakat R (2020). Physical activity and prenatal depression: going beyond statistical significance by assessing the impact of reliable and clinical significant change. *Psychological Medicine* 1–6. https://doi.org/10.1017/S0033291719003714

Received: 30 January 2019
Revised: 9 July 2019
Accepted: 5 September 2019

Author for correspondence:
Marina Vargas-Terrones, E-mail: m.vargas.terrones@gmail.com

Abstract

Background. Previous literature supports exercise as a preventative agent for prenatal depression; however, treatment effects for women at risk for prenatal depression remain unexplored. The purpose of the study was to examine whether exercise can lower depressive symptoms among women who began pregnancy at risk for depression using both a statistical significance and reliable and clinically significant change criteria.

Methods. This study is a secondary analysis of two randomized controlled trials that followed the same exercise protocol. Pregnant women were allocated to an exercise intervention group (IG) or control group (CG). All participants completed the Center for Epidemiological Depression (CES-D) scale at gestational week 9–16 and 36–38. Women with a baseline score ⩾16 were included. A clinically reliable cut-off was calculated as a 7-point change in scores from pre- to post-intervention.

Results. Thirty-six women in the IG and 25 women in the CG scored ⩾16 on the CES-D at baseline. At week 36–38 the IG had a statistically significant lower CES-D score (14.4 ± 8.6) than the CG (19.4 ± 11.1; $p < 0.05$). Twenty-two women in the IG (61%) had a clinically reliable decrease in their post-intervention score compared to eight women in the CG (32%; $p < 0.05$). Among the women who met the reliable change criteria, 18 (81%) in the IG and 7 (88%) in the CG had a score <16 post-intervention, with no difference between groups ($p > 0.05$).

Conclusions. A structured exercise program might be a useful treatment option for women at risk for prenatal depression.

Introduction

Up to 20% of pregnant women experience prenatal depression and ~19% will continue to feel depressive symptoms in the postpartum (Gavin et al., 2005; Woody, Ferrari, Siskind, Whiteford, & Harris, 2017). Prevalence of prenatal depression increases among women who have experienced previous depression, depression in a prior pregnancy, lack of social support, stressful life events, maternal anxiety (Andrews-Fike, 1999), a body mass index (BMI) ⩾30 kg/m^2, lower education, and low socio economic status (Shen, Lin, & Jackson, 2010). Pharmacological treatment options are often prescribed to women with prenatal depression, however many women will choose not to use antidepressants during pregnancy due to the unknown risks medication intake may have on the developing fetus (Latendresse, Elmore, & Deneris, 2017). Untreated prenatal depression can increase the risk for pregnancy complications that impact both the mother and child including preterm birth, intrauterine growth restriction, and early cessation of exclusive breastfeeding (Davalos, Yadon, & Tregellas, 2012; Diego et al., 2009; Figueiredo, Canario, & Field, 2014; Liu, Cnattingius, Bergstrom, Ostberg, & Hjern, 2016).

Recent literature has shown that exercise may be an alternative treatment option for depression during pregnancy (Daley et al., 2015). Guidelines recently released by Canada indicate that pregnant women should aim to achieve 150 min of physical activity every week over at least three sessions (Mottola et al., 2018). Women who meet these guidelines during pregnancy have been shown to have decreased depressive symptoms, tension, fatigue, and anxiety (Gaston

& Prapavessis, 2013). Two randomized controlled trials (RCT) that provided pregnant women with a supervised exercise program three times per week at 16 weeks gestation until delivery, found that fewer women in the intervention group (IG) were at risk for prenatal depression than in the standard care control group (CG) (Perales, Refoyo, Coteron, Bacchi, & Barakat, 2015b; Vargas-Terrones, Barakat, Santacruz, Fernandez-Buhigas, & Mottola, 2018). Similarly, an RCT that provided previously sedentary pregnant women with a walking program found that a brisk walk three times per week for 30 min per session was associated with improved mood and a decreased risk of prenatal depression (Taniguchi & Sato, 2016). Additionally, an RCT that provided pregnant women with aerobic exercise classes once per week for 12 weeks found that the IG had fewer depressive symptoms at the end of the program compared to the standard care CG (El-Rafie, Khafagy, & Gamal, 2016). Finally, a recent meta-analysis found that prenatal exercise reduced the severity of depressive symptoms and the odds of developing prenatal depression by 67% (Davenport et al., 2018).

Depression during pregnancy is often measured by the Centre for Epidemiological Studies-Depression Scale (CES-D) that was developed for diagnosing depression and identifying individuals at risk for depression in the general population (Radloff, 1977). A score ⩾16 on the CES-D has been used as an indicator for depression risk and diagnosis. Research evaluating the effectiveness of exercise during pregnancy on depression scores have mostly included women with a wide range of CES-D scores at baseline, including women who are both considered at risk or have depression at baseline and women who are not at risk (CES-D <16). This reduces the ability to determine if exercise during pregnancy may potentially treat or reduce prenatal depression if women who were not at risk are also a part of the data being assessed.

Furthermore previous exercise and depression literature in pregnant population groups has primarily focused on statistical significance when assessing the impact of exercise on CES-D scores before and after an intervention (Daley et al., 2015; El-Rafie et al., 2016; Perales, Cordero, Vargas Terrones, Lucia, & Barakat, 2015a; Perales et al., 2015b; Taniguchi & Sato, 2016; Vargas-Terrones et al., 2018). To date, no research has used reliable and clinical significant change criteria to identify decreases in depression due to treatment and the clinical likelihood of moving from an 'at risk' to a 'not at risk' depression score following treatment (Jacobson & Truax, 1991). A study that assessed the effectiveness of smoking cessation treatments on depression scores using the CES-D found that using reliable and clinical significance change criteria *v.* statistical significance produced different results and therefore both should be reported (Busch, Wagener, Gregor, Ring, & Borrelli, 2011).

The purpose of the current study was to assess the effectiveness of exercise during pregnancy on CES-D scores among women who have a baseline score ⩾16. In addition to reporting statistical changes in mean values, we calculated a reliable change CES-D index to determine if exercise treatment during pregnancy can achieve this individual level of change among women who are at risk for depression at baseline. Furthermore, if reliable change was shown, clinical significant change through the number of pregnant women who transition from an 'at risk' (e.g. CES-D score ⩾16) to a 'not at risk' (e.g. CES-D score ⩽16) depressive state as a result of treatment would be determined. As previous literature has shown, exercise during pregnancy reduces the prevalence of women who experience depressive symptoms. It is hypothesized that women who receive the exercise intervention will have a greater decrease in CES-D scores than women who receive standard care only. Furthermore, since reliable and clinical significant change criteria to assess the reduction of depression from exercise in pregnant women are unknown, we report the calculated values.

Methods

Study design and recruitment

The present study was a secondary analysis of two RCTs developed by the same research group in Madrid, Spain, and following the same exercise protocol and study methodology. The first RCT (identifier: NCT01696201) was conducted from October 2009 to May 2013 in the University Hospital of Fuenlabrada (Perales et al., 2015b). The second RCT (identifier: NCT02420288) was conducted from October 2014 to December 2016 in the University Hospital of Torrejon (Vargas-Terrones et al., 2018). The research protocol was approved by the Research Ethics Committee of both hospitals and followed the ethical guidelines of the Declaration of Helsinki, which was last modified in 2008.

Information about the studies was given to women with a singleton pregnancy who were <16 weeks pregnant by the attending obstetrician. Women interested in participating contacted the investigators by email or phone, and an information meeting was arranged. According to the exclusion criteria, women with a maternal age <18 or >45 years and women not under medical follow-up throughout pregnancy at the referral hospitals were not included in the studies. Women were also excluded if they had any of the following serious medical conditions: cardiovascular, respiratory or systemic serious disorders, persistent second or third trimester bleeding, placenta previa, ruptured membranes, risk of premature labor, pregnancy-induced hypertension or preeclampsia, and an incompetent cervix. All participants provided signed written informed consent prior to participating in the studies. This study was single-blinded, given that participants were not blinded from the group. The researcher who performed the statistical analysis was blinded. No information about the factors that were studied was given to the participants; however, they had to complete the questionnaires at baseline and at the end of the study.

A randomization by a computer-generated list of numbers was performed with the program EPIDAT 3.1 to allocate the participants into the groups in order of entry: IG and CG. Sixty-one women ($n = 61$) who participated in the previous two studies had a CES-D score that was ⩾16 before the program began. Thirty-six women had been randomized to the IG and participated in the prenatal exercise program and 25 women were in the CG and received standard care only.

Characteristics of the participants

The participants had an initial meeting before gestational week 16 in which they provided data about age, parity, smoking status, education level, occupational activity, previous incidence of miscarriage and current physical activity [frequency, intensity, time, and type (FITT)]. Between weeks 26–28 and 37–39, all women were interviewed regarding the level of physical activity performed during pregnancy (FITT). Medical records were reviewed for pre-pregnancy weight and height to calculate pre-pregnancy BMI. All women who participated in the studies received usual care from

health professionals of the hospitals and the general recommendations of nutrition and exercise. In addition, women who were randomly allocated to the IG participated in a specific exercise program designed for healthy pregnant women.

Exercise intervention (IG)

The exercise intervention program took place in a fitness room inside the hospitals and consisted of three sessions per week from 12–16 gestational weeks to the end of the third trimester (weeks 38–40). In the event of no preterm delivery, 66 to 78 sessions were planned for each participant. To increase program compliance, two to three daily sessions were offered four times a week. The exercise program was designed according to the standards of the American College of Obstetricians and Gynecologists (Gregory, Davies, Mottola, & MacKinnon, 2003), and was similar to previous studies from the same research group (Barakat et al., 2016; Barakat, Pelaez, Montejo, Luaces, & Zakynthinaki, 2011).

Women used a heart rate monitor Polar-FT7 (Polar, Kempele, Finland) to maintain a heart rate intensity of 55% to 60% of heart rate reserve using the Karvonen formula in the aerobic part of the session (Goldberg, Elliot, & Kuehl, 1988). In addition, the Borg Rating of Perceived Exertion (RPE) Scale was used (O'Neill, Cooper, Mills, Boyce, & Hunyor, 1992). Each session consisted of 60 min, distributed as follows: a 10-min-warm-up consisting of 5 min of walking and 5 min of light static stretching of most muscle groups and joint mobility exercises; 25 min of aerobic exercise developed at a moderate intensity through different choreographies; 10 min of muscle strengthening exercises; 5 min of coordination and balance; 5 min of pelvic floor exercises; and at the end of each session, 5–10 min were devoted to stretching and relaxation. The sessions were conducted in groups of 10 to 12 participants and were supervised by a qualified fitness specialist. Extreme stretches, Valsalva maneuver, ballistic movements and jumps were avoided. The exercises performed in the supine position did not exceed two minutes of duration. Adherence to the exercise program was measured by recording attendance at each session.

Usual care (CG)

Women randomly assigned to the CG received general advice from their health care provider about the positive effects of physical activity. Participants in the CG had their usual visits with health care providers during pregnancy, which were equal to the exercise group. Women were not discouraged from exercising. However, women in the CG were asked about their exercise habits once each trimester using a 'Decision Algorithm' (by telephone).

Outcome

Risk of depression was assessed using the Center for Epidemiological Studies-Depression (CES-D) Scale at the beginning of the studies (weeks 12–16) and at the end of the program (gestational weeks 38–39). The CES-D consists of 20 items assessing the different aspects of depressive symptomatology. According to the symptom frequency, each response ranges from 0 (never) to 3 (most days). The score is the sum of the 20 weighted items, and the range of scores is 0–60. If more than four items are missing, the test cannot be considered. A score $\geqslant$16 indicated a risk of depression. This scale is widely distributed and has been used in pregnant populations (El-Rafie et al., 2016; Ko, Yang, & Chiang, 2008; Robledo-Colonia, Sandoval-Restrepo, Mosquera-Valderrama, Escobar-Hurtado, & Ramirez-Velez, 2012). The scale has been translated and validated in Spanish, and it has a high correlation with several scales with a validity between 0.69 and 0.89, a responsiveness of 0.95, a specificity of 0.66 and a reliability of 0.9 (Soler et al., 1997). For the present study, only women at risk of early prenatal depression according to the CES-D questionnaire were included. For this, all the participants from both RCT's with a score $\geqslant$16 at the beginning of the studies were analyzed.

Other included measures

The BMI was calculated by dividing the pre-pregnancy weight (kg) by height (m^2), and women were classified as underweight (BMI < 18.5 kg/m^2), normal weight (BMI $\geqslant$ 18.5 to 24.9 kg/m^2), overweight (BMI $\geqslant$ 25 to 29.9 kg/m^2) or obese (BMI $\geqslant$ 30 kg/m^2). Total gestational weight gain (GWG) was determined by reviewing medical charts and was categorized according to the Institute of Medicine (2009) guidelines for weight gain during pregnancy as excessive or adequate. Excessive GWG was defined as a weight gain of >18.0 kg for underweight, >16.0 kg for normal weight, >11.5 kg for overweight and >9.0 kg for women with obesity (Institute of Medicine and National Research Council Committee to Reexamine, 2009).

Statistical analysis

Descriptive statistics, including means and standard deviations (S.D.) for quantitative variables and percentage for categorical variables, were calculated to examine the maternal characteristics and CES-D scores at baseline. To examine if there were differences between groups at baseline and after the intervention, the χ^2 test and Student's t test were used for categorical and quantitative variables, respectively. One-way ANCOVA was used to calculate post-treatment group differences in CES-D scores. In this analysis baseline CES-D scores served as a covariate. The study analyses were performed using the Statistical Package for Social Sciences data software, version 20.0.

To calculate the threshold for reliable change, the equation employed by Busch et al., was used (Busch et al., 2011). This equation takes into account the internal consistency of the CES-D scale and the standard deviation for the population in the current study. The threshold for reliable change is 1.96 times the standard error of difference between the two-time points (baseline and end of intervention). To calculate the standard error of difference the formula provided by Jacobson & Truax (1991) was used (Fig. 1) (Busch et al., 2011). The standard deviation at baseline is referred to as SDpre, and in the current population this was 5.14. The Cronbach's α (reliability of the measure) was 0.752. A standard error of difference was calculated as 3.6 and therefore the reliable change cut-off for this group was 7.09 (3.6 multiplied by 1.96), rounded to 7. χ^2 analysis was used to determine how many women from each treatment condition represented this reliable change decrease in CES-D from baseline. To shed light on clinical significance, χ^2 analysis was used to determine how many women in each treatment group who met the reliable change decrease in CES-D from baseline also transitioned from an 'at risk' (equal or above 16 on the CES-D) to 'not at risk' (below 16 on the CES-D) scale range.

$$\sqrt{2(SD_{Pre}\sqrt{1-\alpha)^2}}$$

Fig. 1. Equation for the calculation of reliable change (Jacobson & Truax, 1991).

Results

Among the 61 pregnant women who had a baseline CES-D score ⩾16 there were no differences observed between the two groups ($p > 0.05$) for age, parity, pre-pregnancy BMI, pre-pregnancy weight, smoking status before and during pregnancy, occupational activity level, education level, previous incidence of miscarriage, total GWG, and number of women who exceeded GWG recommendations (Table 1). Examining baseline characteristics from the two studies, there were no differences between the participants from both studies in any of the above mentioned characteristics ($p > 0.05$), except on the education level, which was higher in the participants from the first study (Hospital of Fuenlabrada) than the second study (Hospital of Torrejon) (Elementary school 34.2% *v.* 65.2%, High school 39.5% *v.* 34.8%, University 26.3% *v.* 0.0% respectively; $p < 0.05$).

Regarding the exercise habits of the participant during pregnancy, no significant differences were found in the level of physical activity during pregnancy between groups, excluding the exercise developed in the intervention program ($p > 0.05$).

There was no difference between the groups for their initial CES-D score ($t_{59} = 0.870$; $p > 0.05$). Controlling for initial CES-D scores, the IG had significantly lower post-treatment CES-D scores than the CG ($\eta^2 = 0.054$; $F_{1,58} = 4.790$; $p < 0.05$), with a five-point difference between groups (IG = 14.4 ± 8.6 *v.* CG = 19.4 ± 11.1) (Table 2). There were no differences between the two studies in the baseline and the post-treatment CES-D score ($p > 0.05$), neither in the percentage of women at risk of depression at the end of the intervention ($p > 0.05$).

When analyzing the percentage of women who had a post-intervention decrease in their CES-D score, there was a significant difference between groups, with 83.3% ($n = 30$) of women in the IG who decreased their score compared to 56.0% in the CG ($n = 14$; $\chi^2 = 5.483$; $p < 0.05$). Furthermore, more women in the IG met the reliable change threshold cut off (⩾7 point decrease in score) compared to the CG (IG: $n = 22$, 61%; CG: $n = 8$, 32%; $\chi^2 = 5.003$, $p < 0.05$). Of the 22 women in the IG who met the reliable change criteria, 18 (81%) had a score <16 post-intervention. Of the eight women in the CG who met the reliable change criteria, 7 (88%) had a score <16, with no difference between both groups ($\chi^2 = 2.283$, $p > 0.05$). When using the criteria of how many women crossed the ⩾16 at risk cut-off to now be <16 on the CES-D, there was no significant difference ($\chi^2 = 0.338$; $p > 0.05$) between groups (IG-55.6%; $n = 20$ *v.* CG-48.0%; $n = 12$).

Table 1. Maternal characteristics of the IG and CG

	IG ($n = 36$)	CG ($n = 25$)
Maternal characteristic at the beginning of the study		
Maternal age (years)	32.5 ± 3.3	32.6 ± 4.7
BMI (kg/m^2)	23.5 ± 3.7	24.2 ± 6.1
BMI categories, *n*/%		
Underweight	1/2.7	2/8.0
Normal weight	24/66.6	16/64.0
Overweight	9/25.0	3/12.0
Obese	2/5.6	4/16.0
Pre-pregnancy weight (kg)	62.5 ± 10.5	65.2 ± 18.2
Occupational activity, *n*/%		
Unemployed/Homemaker	17/47.2	8/32.0
Sedentary job	12/33.3	7/28.0
Active job	7/19.4	10/40.0
Level of education, *n*/%		
Elementary school	17/47.2	11/44.0
High school/College	15/41.7	8/32.0
University	4/11.1	6/24.0
Parity, *n*/%		
Nulliparous	21/58.3	13/52.0
Primiparous	14/38.8	11/44.0
Multiparous	1/2.8	1/4.0
Smoking before pregnancy, *n*/%		
Yes	15/41.7	9/36.0
No	21/58.3	16/64.0
Smoking during pregnancy, *n*/%		
Yes	3/8.3	6/24.0
No	33/91.7	19/76.0
Previous miscarriage, *n*/%		
None	28/77.8	18/72.0
1	8/22.2	6/24.0
2	0/0.0	1/4.0
Maternal characteristics post-intervention		
Total GWG	12.4 ± 4.5	13.3 ± 4.8
GWG categories, *n*/%		
Adequate	24/66.7	14/56.0
Excessive	12/33.3	10/40.0

BMI, body mass index.
There were no statistical differences between groups at baseline ($p > 0.05$).
Data are expressed as mean ± s.d. unless otherwise indicated.

Discussion

The current study examined the effect of physical exercise during pregnancy on the incidence of prenatal depression according to the CES-D scale among women who were 'at risk' of depression at the beginning of pregnancy (initial CES-D score ⩾16). More specifically, this study examined whether exercise can lower CES-D scores among women who began pregnancy at risk for depression. To this end, both a statistical significance and reliable and clinically significant change criteria were used.

With respect to statistical significance, we found evidence for a moderate size effect where IG participants had lower post-treatment CES-D scores compared to their CG counterparts, with an average difference of five points between groups. It was observed that the CES-D mean score decreased below at risk levels (<16) in the IG, whereas the mean in the CG remained above 16.

Table 2. CES-D scores at the beginning and at the end of the study for each group

CES-D score for the whole group ($n = 61$)	Intervention $n = 36$	Control $n = 25$	p
CES-D score at baseline	23.2 ± 5.2	22.0 ± 5.0	0.39
CES-D score at week 38	14.4 ± 8.6	19.4 ± 11.1	0.03

CES-D, Center for Epidemiological Studies-Depression Scale.

With respect to a reliable change, more women in the IG (22; 61%) compared to women in the CG (8; 32%) decreased post intervention CES-D scores by the 7 point threshold for reliable change, which is in accordance with the results that show that participants of the IG had a higher percentage of women who reduced their CES-D score than among women of CG. With respect to clinical significance, for those women who met the 7 point threshold for reliable change there was no difference in the number of women who had a score <16 post intervention in the CG and IG. Specifically, 7 out of the 8 women in the CG (88%) who met the reliable change threshold went below the 16 point cut-off post intervention compared to 18 out of the 22 women in the IG (81%). These findings suggest that irrespective of treatment, pregnant women will benefit proportionally in a high manner with respect to transitioning from 'being at risk' (CES-D ⩾16) to being 'not at risk' (CES-D <16) of clinical depression when the threshold for reliable change is achieved. Research in non-pregnant populations suggests that although it is beneficial to see a CES-D score below 16 (not at risk for depression), a clinically significant change based on a reliable change threshold may be more telling of decreased depression risk in the future. This is important in the pregnant population group as post-delivery women are at risk of postpartum depression and this risk is higher among women who experienced depressive symptoms during their pregnancy.

To our knowledge no other study has examined the effects of an exercise intervention during pregnancy among women who have an initial CES-D score ⩾16 (at risk of depression). Previous exercise interventions have included both women at risk and not at risk of depression in the same sample (Robertson, Grace, Wallington, & Stewart, 2004). This study was able to specifically highlight the potential treatment effects exercise may have on prenatal depression and depressive symptoms. Not surprising, including women at risk of depression only, the current study saw a greater decrease in mean CES-D scores compared to studies that included both at risk and not-at-risk women. One study that provided a 3-month long exercise intervention saw that the intervention group on average had a 4-point decrease post-intervention (Robledo-Colonia et al., 2012). Similarly, a 12 week long exercise program showed a just over 5 point decrease in scores post-intervention (El-Rafie et al., 2016). In the current study the IG had an 8 point decrease in scores. Interestingly, only those in the IG reduced their CES-D mean change scores beyond the reliable change threshold (>7 point decrease). These results are promising as they show that exercise may have a stronger treatment effect among women who are at risk for depression. A trend that was found in our study and is consistent with the previously mentioned RCTs is the fact that the standard care CG did not have a pre-post CES-D change score, and their mean scores remained >16 post-intervention. This study therefore adds to the existing evidence that providing no intervention during pregnancy will not improve depressive symptoms with those with an initial CES-D score above 16.

Strengths of the current study include the inclusion of data from two RCTs that included women with similar baseline CES-D scores and women who had high adherence to the exercise intervention [>80% attendance to exercise classes (Perales et al., 2015b; Vargas-Terrones et al., 2018]. This is the first study to specifically assess women with an initial CES-D score ⩾16 only, therefore isolating and highlighting the results specifically for women who are experiencing early prenatal depressive symptoms and are at risk for depression. Finally, this is the first study to calculate and assess a population specific reliable change index to detect a potential clinically significant change in CES-D scores which was achieved by more women in the IG. A limitation of the current study is that it utilized a convenience sample from the two RCTs, hence *a priori* sample size and power calculations were not conducted. Future studies should include adequately powered RCTs that focus exclusively on women at risk of depression during pregnancy (⩾16). In addition, future studies may want to determine whether these findings can be replicated with pregnant women with higher baseline CES-D scores, indicating more severe symptoms of depression. Furthermore, as prenatal depression is indicative of postnatal depression, future studies should include postpartum follow up to assess the impact of prenatal exercise on the risk of postnatal depression. In addition, it would be advisable to develop studies that include information about clinical diagnosis of depression and to add a follow up of factors (i.e. social support) that may influence depression symptoms during the intervention. Finally, future RCTs should include an attention CG in addition or instead of a CG to disentangle non-specific *v.* specific (exercise) depression treatment effects.

In conclusion, we provide both statistical as well as reliable and clinical significance change evidence that structured exercise during pregnancy reduced depressive symptoms, thus not only can exercise be a preventive agent, but also a potentially feasible treatment option for prenatal depression. Exercise during pregnancy may be an effective way to promote psychological well-being during pregnancy and this can improve the overall health of both mother and fetus.

Acknowledgements. The authors would like to acknowledge the technical assistance of the Gynecology and Obstetrics Service of Hospital de Torrejon and Hospital de Fuenlabrada (Madrid) and the collaboration of all the participants.

Financial support. None.

Conflict of interest. None.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. The research protocol was approved by the Research Ethics Committee of the hospitals.

References

Andrews-Fike, C. (1999). A review of postpartum depression. *Primary Care Companion to the Journal of Clinical Psychiatry*, *1*, 9–14.

Barakat, R., Pelaez, M., Montejo, R., Luaces, M., & Zakynthinaki, M. (2011). Exercise during pregnancy improves maternal health perception: A randomized controlled trial. *American Journal of Obstetrics and Gynecology*, *204*, 402.e1–402.e7.

Barakat, R., Pelaez, M., Cordero, Y., Perales, M., Lopez, C., Coteron, J., & Mottola, M. F. (2016). Exercise during pregnancy protects against hypertension and macrosomia: Randomized clinical trial. *American Journal of Obstetrics and Gynecology*, *214*, 649.e1–649.e8.

Busch, A. M., Wagener, T. L., Gregor, K. L., Ring, K. T., & Borrelli, B. (2011). Utilizing reliable and clinically significant change criteria to assess for the development of depression during smoking cessation treatment: The importance of tracking idiographic change. *Addictive Behaviors*, *36*, 1228–1232.

Daley, A. J., Foster, L., Long, G., Palmer, C., Robinson, O., Walmsley, H., & Ward, R. (2015). The effectiveness of exercise for the prevention and treatment of antenatal depression: Systematic review with meta-analysis. *BJOG: An International Journal of Obstetrics and Gynaecology*, *122*, 57–62.

Davalos, D. B., Yadon, C. A., & Tregellas, H. C. (2012). Untreated prenatal maternal depression and the potential risks to offspring: A review. *Archives of Women's Mental Health*, *15*, 1–14.

Davenport, M. H., McCurdy, A. P., Mottola, M. F., Skow, R. J., Meah, V. L., Poitras, V. J., … Ruchat, S.-M. (2018). Impact of prenatal exercise on both prenatal and postnatal anxiety and depressive symptoms: A systematic review and meta-analysis. *British Journal of Sports Medicine*, *52*, 1376–1385.

Diego, M. A., Field, T., Hernandez-Reif, M., Schanberg, S., Kuhn, C., & Gonzalez-Quintero, V. H. (2009). Prenatal depression restricts fetal growth. *Early Human Development*, *85*, 65–70.

El-Rafie, M. M., Khafagy, G. M., & Gamal, M. G. (2016). Effect of aerobic exercise during pregnancy on antenatal depression. *International Journal of Women's Health*, *8*, 53–57.

Figueiredo, B., Canario, C., & Field, T. (2014). Breastfeeding is negatively affected by prenatal depression and reduces postpartum depression. *Psychological Medicine*, *44*, 927–936.

Gaston, A., & Prapavessis, H. (2013). Tired, moody and pregnant? Exercise may be the answer. *Psychology & Health*, *28*, 1353–1369.

Gavin, N. I., Gaynes, B. N., Lohr, K. N., Meltzer-Brody, S., Gartlehner, G., & Swinson, T. (2005). Perinatal depression: A systematic review of prevalence and incidence. *Obstetrics & Gynecology*, *106*, 1071–1083.

Goldberg, L., Elliot, D. L., & Kuehl, K. S. (1988). Assessment of exercise intensity formulas by use of ventilatory threshold. *Chest*, *94*, 95–98.

Gregory A. L. Davies, L. A. W., Mottola M. F., & MacKinnon C. (2003). Joint SOGC/CSEP clinical practice guideline: Exercise in pregnancy and the postpartum period. *Canadian Journal of Applied Physiology*, *28*, 330–341.

Institute of Medicine & National Research Council Committee to Reexamine IOM Pregnancy Weight Guidelines. (2009). The national academies collection: Reports funded by national institutes of health. In K. M. Rasmussen & A. L. Yaktine (Eds.), *Weight gain during pregnancy: Reexamining the guidelines*. Washington, DC: National Academies Press. Summary, Table S-1. Available from: https://www.ncbi.nlm.nih.gov/books/NBK32799/.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19.

Ko, Y. L., Yang, C. L., & Chiang, L. C. (2008). Effects of postpartum exercise program on fatigue and depression during 'doing-the-month' period. *The Journal of Nursing Research: JNR*, *16*, 177–186.

Latendresse, G., Elmore, C., & Deneris, A. (2017). Selective serotonin reuptake inhibitors as first-line antidepressant therapy for perinatal depression. *Journal of Midwifery & Women's Health*, *62*, 317–328.

Liu, C., Cnattingius, S., Bergstrom, M., Ostberg, V., & Hjern, A. (2016). Prenatal parental depression and preterm birth: A national cohort study. *BJOG: An International Journal of Obstetrics and Gynaecology*, *123*, 1973–1982.

Mottola, M. F., Davenport, M. H., Ruchat, S.-M., Davies, G. A., Poitras, V. J., Gray, C. E., … Zehr, L. (2018). 2019 Canadian guideline for physical activity throughout pregnancy. *British Journal of Sports Medicine*, *52*, 1339–1346.

O'Neill, M. E., Cooper, K. A., Mills, C. M., Boyce, E. S., & Hunyor, S. N. (1992). Accuracy of Borg's ratings of perceived exertion in the prediction of heart rates during pregnancy. *British Journal of Sports Medicine*, *26*, 121–124.

Perales, M., Cordero, Y., Vargas Terrones, M., Lucia, A., & Barakat, R. (2015a). Exercise and depression in overweight and obese pregnant women: A randomised controlled trial. *Archivos de Medicina del Deporte*, *167*, 70–77.

Perales, M., Refoyo, I., Coteron, J., Bacchi, M., & Barakat, R. (2015b). Exercise during pregnancy attenuates prenatal depression: A randomized controlled trial. *Evaluation & The Health Professions*, *38*, 59–72.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401.

Robertson, E., Grace, S., Wallington, T., & Stewart, D. E. (2004). Antenatal risk factors for postpartum depression: A synthesis of recent literature. *General Hospital Psychiatry*, *26*, 289–295.

Robledo-Colonia, A. F., Sandoval-Restrepo, N., Mosquera-Valderrama, Y. F., Escobar-Hurtado, C., & Ramirez-Velez, R. (2012). Aerobic exercise training during pregnancy reduces depressive symptoms in nulliparous women: A randomised trial. *Journal of Physiotherapy*, *58*, 9–15.

Shen, J. J., Lin, F., & Jackson, T. (2010). Risk of prenatal depression: Differences by race. *Ethnicity & Disease*, *20*, 35–39.

Soler, J., Perez-Sola, V., Puigdemont, D., Perez-Blanco, J., Figueres, M., & Alvarez, E. (1997). [Validation study of the center for epidemiological studies-depression of a Spanish population of patients with affective disorders]. *Actas luso-espanolas de neurologia, psiquiatria y ciencias afines*, *25*, 243–249.

Taniguchi, C., & Sato, C. (2016). Home-based walking during pregnancy affects mood and birth outcomes among sedentary women: A randomized controlled trial. *International Journal of Nursing Practice*, *22*, 420–426.

Vargas-Terrones, M., Barakat, R., Santacruz, B., Fernandez-Buhigas, I., & Mottola, M. F. (2018). Physical exercise programme during pregnancy decreases perinatal depression risk: A randomised controlled trial. *British Journal of Sports Medicine*, *53*, 348–353.

Woody, C. A., Ferrari, A. J., Siskind, D. J., Whiteford, H. A., & Harris, M. G. (2017). A systematic review and meta-regression of the prevalence and incidence of perinatal depression. *Journal of Affective Disorders*, *219*, 86–92.