A pragmatic randomized controlled trial to evaluate the effectiveness of a facilitated exercise intervention as a treatment for postnatal depression: the PAM-PeRS trial

Daley, Amanda; Blamey, R. V.; Jolly, Catherine; Roalfe, A. K.; Turner, K. M.; Coleman, Sarah; Mcguinness, M.; Jones, I.; Sharp, D. J.; Macarthur, C.

DOI:
10.1017/S0033291715000409

A pragmatic randomised controlled trial to evaluate the effectiveness of a facilitated exercise intervention as a treatment for postnatal depression: the PAM-PeRS trial

AJ Daley[1], PhD, senior lecturer in health psychology, a.daley@bham.ac.uk
RV Blamey[1], BMed Sc research fellow, r.v.blamey@bham.ac.uk
K Jolly[2], PhD, professor of public health and primary care, c.b.jolly@bham.ac.uk
AK Roalfe[1], MSc, senior lecturer in medical statistics, a.k.roalfe@bham.ac.uk
KM Turner[3], PhD, senior lecturer in primary care, katrina.turner@bristol.ac.uk
S Coleman[1], BA Hons, research associate, s.j.coleman.1@bham.ac.uk
M McGuinness[4], PhD, consultant clinical psychologist, Mary.McGuinness@bsmhft.nhs.uk
I Jones[5], MRCPsych, PhD, professor of perinatal psychiatry, jonesir1@cf.ac.uk
DJ Sharp[3], FRCGP, PhD, professor of primary care, Debbie.Sharp@bristol.ac.uk
C MacArthur[2], PhD, professor of maternal and child epidemiology, c.macarthur@bham.ac.uk

[1]Primary Care Clinical Sciences, School of Health and Population Sciences, University of Birmingham, UK

[2]Public Health, Epidemiology and Biostatics, School of Health and Population Sciences, University of Birmingham, UK

[3]Academic Unit of Primary Health Care, University of Bristol, Bristol, UK

[4]Perinatal Mental Health Service, The Barberry, Birmingham and Solihull Mental Health Foundation Trust, Birmingham, UK

[5]Institute of Psychological Medicine & Clinical Neurology, Cardiff University, Cardiff, Wales, UK

Corresponding author
Professor Kate Jolly
Public Health, Epidemiology and Biostatics
School of Health and Population Sciences
University of Birmingham, UK
0121 4147552
c.b.jolly@bham.ac.uk

Word count: 4230

Number of Tables:4; number of figures 1; supplementary tables 2

Abstract

Background: Postnatal depression affects about 10-15% of women in the year after giving birth. Many women and healthcare professionals would like an effective and accessible non pharmacological treatment for postnatal depression.

Methods: Women who fulfilled the ICD-10 criteria for major depression in the first six months postnatally were randomised to receive usual care plus a facilitated exercise intervention or usual care only. The intervention involved two face to face consultations and two telephone support calls with a physical activity facilitator over six months to support participants to engage in regular exercise. The primary outcome was symptoms of depression using the Edinburgh Postnatal Depression Scale (EPDS) at six month post-randomisation. Secondary outcomes included EPDS score as a binary variable (recovered and improved) at six and 12 month post-randomisation.

Results: 146 women were potentially eligible and 94 were randomised. 34% reported thoughts of self-harming at baseline. After adjusting for baseline EPDS, analyses revealed a -2.04 mean difference in EPDS score, favouring the exercise group (95% CI: -4.11 to 0.03, p=0.05). When also adjusting for pre-specified demographic variables the effect was larger and statistically significant (mean difference=-2.26, 95% CI:-4.36 to -0.16, p=0.03). Based on EPDS score a larger proportion of the intervention group were recovered (46.5% versus 23.8%, p=0.03) compared with usual care at six months follow-up.

Conclusions: This trial shows an exercise intervention that involved encouragement to exercise and to seek out social support to exercise may be an effective treatment for women with postnatal depression, including those with thoughts of self-harming.

Keywords: exercise, postnatal, postpartum, depression, primary care

Introduction

Postnatal depression affects about 10-15% of women in the year after birth (O'Hara et al., 1996; Gaynes et al., 2005) with symptoms including low mood, fatigue, anxiety, thoughts of self harming and poor mother-infant interactions (Beck, 1992, 1995; Murray, 1992; Cooper et al., 1995; American Psychiatric Association, 2000). Current treatments recommended for postnatal depression include antidepressants and psychological therapies (National Institute for Health and Care Excellence [NICE], 2007). There are some risks in taking antidepressants and reluctance amongst new mothers to take them (Turner et al., 2008; Turner et al., 2010). Although non pharmacological interventions can be effective (Dennis et al., 2007; Morrell et al., 2009; Sharp et al., 2010) there can be long waiting lists to access treatment and they can be costly. Exercise has the potential to address barriers associated with traditional treatments: it is free, accessible and is without stigma or side effects (Daley et al., 2007; Lewis et al., 2014).

The American College of Obstetrics and Gynecology, (2002) and NICE (2007) in England have stated that self help strategies such as exercise should be considered as a treatment for postnatal depression, although this guidance from NICE was based on evidence at the time from trials that had recruited general populations and two small trials of women at risk of postnatal depression. A systematic review with meta-analysis by the trial authors found five small trials in women at risk of postnatal depression, but no samples with *clinically* defined depression (Daley et al., 2009). The review showed that exercise significantly reduced postnatal depression scores relative to comparators, but significant heterogeneity was present and the effect size became non-significant when the trial that included exercise with social support was excluded (Armstrong et al., 2003). We recently updated this review (Blamey et al., 2012) and no further trials that had recruited women with clinically diagnosed depression have been published.

It remains uncertain therefore whether exercise reduces symptoms of postnatal depression and a high quality trial in women with clinically diagnosed depression was considered essential. Following a pilot trial this RCT evaluated the effectiveness of a facilitated exercise intervention as a treatment for postnatal depression alongside any usual care compared with a group who only received usual care (Daley et al., 2008). The study was purposively designed as a pragmatic trial where women were eligible for inclusion regardless of whether they were receiving or not receiving other specific treatments (i.e. antidepressants or psychological intervention) because it would not have been safe or ethical to interrupt or withdraw usual care and test exercise alone until there is evidence that exercise is likely to provide benefit in this population. We hypothesised that participants receiving the exercise intervention would report lower depressive symptoms scores at follow up than those receiving usual care.

Design and overview

Setting and participants

147 family medicine practices in Birmingham, UK were invited to assist with recruitment, 67 agreed. Recruitment commenced in April 2010 and follow up completed in April 2013. Women were eligible if they were within six months of giving birth, aged 18 years or more and had an ICD-10 diagnosis of a major depressive episode (WHO, 2011), following initial screening using the Edinburgh Postnatal Depression Scale (EPDS) (Cox et al., 1987) and a clinical diagnostic interview (Lewis et al., 1992). Women with a diagnosis of mixed anxiety and depression were also eligible. Patients were excluded if they were pregnant again, experiencing psychotic symptoms or dependent on illicit drugs or alcohol. Women needed to be currently inactive (not meeting the current guidelines for physical activity) (DOH, 2011).

Patient identification and diagnosis of depression

Women who had recently given birth and who lived within two primary care trusts in Birmingham were identified from the Child Health System (CHS), a centralised computer registration system of all new births in the areas which allowed systematic identification of women in the 67 participating family medicine practices. Every two weeks the CHS completed search lists of new births in participating practices and mailed the trial information letter and EPDS to these women. Women were sent this at 10-14 weeks after birth and there were several reasons for this; many women feel overwhelmed when they have a new baby to care for and may be less inclined to participate in a research study in the first few weeks of giving birth. Having a contact date of 10-14 weeks, rather than shortly after giving birth, gave women the chance to get into a routine with their baby before committing to a long term research study. Symptoms of depression in some cases can resolve naturally after a period of adjustment to the baby and this study was focused on identifying those women who would need assistance with resolving their depression. Women who scored 10+ on this first EPDS (EPDS-1) then completed a second EPDS (EPDS-2) two weeks later by telephone to rule out the possibility of transient depression. Women who scored 13+ on EPDS-2 then completed the Clinical Interview Schedule-Revised (CIS-R) (Lewis, 1994; Jenkins et al., 1997) at a home visit to ensure they met the ICD-10 criteria for experiencing a major depressive episode.

Other recruitment methods

Women who presented with depression after six weeks, but within six months of giving birth were also able to participate via referral from their GP or health visitor. Health professionals from two perinatal mental health community services were also able to refer women. Women identified in these ways also had to meet all the study entry criteria as previously described.

Outcomes, process measures and allocation of participants to trial groups

The primary outcome was difference in mean EPDS score (adjusted for baseline) between the groups at six month follow-up. We used the EPDS as our primary outcome of symptoms of depression so that we could compare the results with other postnatal depression treatment trials as the use of the EPDS is standard practice within postnatal depression research. It can be very difficult to obtain high follow up in women with postnatal depression and using the EPDS, rather than a clinical interview, was considered to provide the best opportunity to achieve high follow up rates.

Secondary outcomes were difference in mean EPDS score at 12 month follow up and rates of improved and recovered responses on the EPDS at follow up. Participants completed the SF-12 (Ware et al., 1996), EQ-5D (Brooks, 1996), a body image subscale (Kumar et al., 1984) the subjective vitality scale (Bostic et al., 2000) and items relating to perceived social support (Sallis et al., 1987; Singleton et al., 2003) and self-efficacy for exercise (Sallis et al., 1988) were also completed. Body weight and height were measured objectively. Physical activity was measured by self report (IPAQ-short) (Craig et al., 2003) and objectively using the Actiheart device.

During intervention weeks 4, 8, 12, 16, 20 the intervention group were asked to complete exercise diary logs as previously used in our pilot trial (Daley et al., 2008). Logs were completed with the physical activity facilitator (PAF) in person, over the telephone or sent/returned by post.

Randomisation, concealment and blinding

An internet randomisation service was used to allocate participants to the trial groups and was concealed from researchers involved in recruiting and randomising participants. The allocation sequence was generated using a computer programme with random permuted blocks of size 10. Participants were randomised into one of two trial groups (50:50 split) and further randomised for wearing the Actiheart or not (60:40 split). The randomisation was

stratified by EPDS score (13-16 and 17+). The person delivering the intervention was not involved in recruiting participants. Participants, researchers and those delivering the intervention could not be blinded to group allocation. The primary outcome, EPDS score, was completed by postal questionnaire (along with other secondary outcomes).

Data collection

Eligible women were asked to provide written informed consent. At baseline and six and 12 month post randomisation follow up, the questionnaires were mailed to participants and collected at home visits by the research team or returned by post. Those randomised to wear the Actiheart had the device fitted at the home visit. The research team collected the Actiheart from these participants and the IPAQ was then completed. Participants who had worn the Actiheart at baseline were asked to wear it again at follow up. Participants were asked to wear it the Actiheart for three consecutive days and to then complete the IPAQ.

Intervention

A detailed description of the six month intervention can be found in the published protocol (Daley et al., 2012). The initial goal (weeks 1-12) was for participants to progress towards accumulating 30 minutes of moderate intensity exercise on three days per week. During weeks 13-24 participants were encouraged to work towards accumulating 30 minutes of moderate intensity exercise on 3-5 days per week. Similar to our pilot trial (Daley et al., 2008) the intervention involved two face to face personalised exercise consultations (during months 1 and 2) and telephone calls (during months 3 and 4). The face to face consultations were centred on equipping women with the skills, knowledge and confidence needed to participate in regular exercise and were delivered by a PAF in participants homes. Consultations lasted between 40-60 min. Participants were given a pedometer. Four weeks later participants had a second consultation centred on the prevention of relapse back to an inactive lifestyle and/or improving maintenance of an active lifestyle. Telephone support

calls (15-20 minutes) were made during months three and four of the intervention. Information leaflets were mailed in months three, four, five and six of the intervention to further encourage exercise participation.

Usual care comparator

Usual care could have included women spontaneously consulting their GP and given active treatment or just consultation to discuss symptoms, or informal counselling from their health visitor or referral by their health visitor to the GP, or that they consulted no one and had no treatment**.** The usual care group were sent the study "*Looking after yourself*" leaflet at baseline and exercise was not further encouraged beyond receipt of this single leaflet.

Sample size and statistical analyses

A sample size of 83 patients randomised to each group (n=166) would be sufficient to detect a 1.95 (SD =3.9, Heh et al., 2008) unit difference in EPDS score between the groups at the six month follow-up with 90% power, 5% significance level. This increased to 104 in each arm when allowance was made for a 20% dropout rate at six month follow-up (n=208). When considered as a continuous variable two EPDS points is considered a moderate effect size change (0.5 SD) (Matthey, 2004).

All analyses used the intention to treat principle, whereby participants were analysed in the group to which they were randomised. All participants that had complete data for the EPDS at 6 months were included in the primary analysis. The primary analysis compared mean EPDS score at six month post randomisation follow-up between trial groups, adjusting for baseline scores using analysis of covariance. EPDS-2 score was used as the baseline score. Similar analyses were conducted for the secondary outcomes (except physical activity) adjusting for baseline outcome score and baseline EPDS. Secondary analyses adjusted for baseline score, baseline EPDS and covariates (age, baseline weight and ethnicity) again using analysis of covariance. To explore longer term effects, a repeated measures mixed model

analysis of the primary and secondary outcomes was undertaken, comparing groups across the follow ups. Adjustments were made for covariates as previously described. Bootstrapped analyses (1000 repetitions) were performed where residuals were non-normally distributed.

To assess clinically meaningful change on the EPDS following the principles of Jacobson and Truax (1991), analysis of the proportion of women per group who were 'recovered' (i.e. a reduction in EPDS score of four points or more plus scored less than 13 points) versus not recovered (Matthey, 2004) was conducted using Chi-squared tests. Categorising data in this way provides a better indication of whether real changes have occurred at the level of the individual participant (refer to Jacobson and Truax 1991 and Matthey 2004 for a more detailed discussion). In addition, in order that we could compare our findings with other previous postnatal depression treatment trials we compared those who improved (i.e. proportions below 13 on the EPDS) versus not improved.

The IPAQ results were negatively skewed and comparison between the groups was undertaken using quantile regression analysis, with baseline physical activity included as a covariate. All statistical analysis was carried out using SAS version 9.2 and Stata version 12.

Results

Study invitations and EPDS-1 were sent by CHS to 9983 women and 1068 were returned (10.7%). An additional 82 women recruited via other methods completed EPDS-1. Of these, 436/1150 (37.9%) scored 10+ on EPDS-1, of whom 146/436 (33.5%) scored 13+ on EPDS-2 two weeks later and were offered a home visit to diagnose depression using the CIS-R and to assess other eligibility criteria. Of these 146 women, 100 (68.5%) were fully eligible and 94 were randomised (Figure 1). The characteristics of responders and non-responders to EPDS-1 were similar in terms of age, deprivation (McLennan et al., 2011) and ethnicity.

Randomisation and participants' baseline characteristics

Our sample size calculation indicated that 166 participants needed to be recruited, rising to 208 with 20% loss to follow up. We did not meet this target and 94 participants were randomised to the exercise plus usual care group (n=47) or the usual care only group (n=47). Table 1 describes the baseline characteristics which shows a good balance between groups on these variables, although more usual care participants were receiving counselling/psychological support than the intervention group (21% versus 7%). The majority of participants lived in the two highest deprivation quartiles (74/94, 79%) and 37% (35/94) were of non-white ethnicity. The questionnaire assessments (EPDS and/or CIS-R) showed that a large proportion of participants had thoughts of self-harming (32/94, 34%). 17 (18.1%) participants were diagnosed on the CIS-R with a severe depressive episode, 50 (53.2%) moderate severe episode, 15 (15.9%) mild episode and 12 (12.8%) were diagnosed with mixed anxiety and depressive disorder. These diagnoses were balanced across groups.

Information about follow up

Rates of breastfeeding, risk of self harm, use of antidepressants and counselling (and both treatments) at six and 12 month post randomisation (supplementary Table 1). More usual care participants were also receiving psychological support at follow up than in the exercise group.

At six months post randomisation 85/94 (90.4%) participants completed follow up for the primary outcome and at 12 months post randomisation the follow up rate was 79/94 (84.0%). Participants lost to follow up tended to be marginally younger compared to those who were followed up but were similar in terms of group allocation, baseline mean EPDS score, deprivation quartile score (McLennan et al., 2010) and ethnicity.

Primary outcome

After adjusting for baseline scores, analyses revealed a -2.04 point mean difference in EPDS score , favouring the exercise intervention group (95% CI: -4.11 to 0.03, p=0.053).

When adjusting for baseline EPDS score and pre-specified demographic variables, the mean difference increased and was statistically significant (-2.26, 95% CI, -4.36 to -0.16, p=0.03). Tables 2 and 3.

Secondary outcomes

At six months significantly more of the intervention group were considered 'recovered' and no longer a case (46.5% versus 23.8%, p=0.03) compared to usual care. By 12 months, the proportions considered 'recovered' (i.e. a reduction in EPDS score of four points or more plus scored less than 13 points) had increased in both groups (51.2% versus 36.8%) but the difference between the groups was reduced and not statistically significant. Table 4.

The exercise group had significantly higher social support scores than usual care at six months in both adjusted models however the effect was not sustained at 12 months. In the fully adjusted analysis there was some evidence of a difference in vitality scores at 6 months, favouring the exercise group (p=0.054), but this effect was not significant at 12 months (p=0.09). The exercise group reported significantly lower social support at 12 months for exercise–family rewards scores (p=0.04) than usual care when adjusting for baseline score, but there was no significant difference between the groups in the fully adjusted model (Tables 2 and 3).

There were no significant differences between the groups on any of the physical activity outcomes derived from the IPAQ (supplementary table 2). Only eight participants (8.5%) wore the Actiheart and provided useable data for analyses at both six and 12 months follow up. These data are therefore not reported since they do not provide reliable information on which to base any conclusions (see discussion later) but it is available from the first author on request.

Intervention implementation and adherence to the intervention goals

Delivery of the various intervention components was very high; 41/47 (87%) participants received all four individual contacts (two consultations and two telephone support calls, 43/47 (91.4%), 44/47 (93.6%) and 46/47 (97.9%) received at least three, two and one respectively of the individual intervention contacts.

A total of 163/235 (69.4%) (i.e. 47 participants x 5 logs) of the exercise logs were completed by the intervention group, with the majority (40/47) completing at least 3/5 logs. Only four participants failed to complete a single log. In the logs participants reported completing an overall average of 146 (SD=142) min of moderate/vigorous exercise. The exercise group reported increased levels of exercise over the course of the intervention; in logs one to five participants reported means per week of 161.1, 217.5, 203.2, 222.5 and 245.0 min of moderate/vigorous exercise respectively (range equates to 23-35 min per day). The most common types of exercise reported were brisk walking with pram, brisk walking without pram, exercise DVDs, Wii fit workouts, jogging and swimming.

Discussion

Women diagnosed with postnatal depression and randomised to a facilitated home based exercise intervention for six months reported lower mean EPDS scores than those randomised to usual care only. Depressed women were twice as likely to report a clinically meaningful change in their EPDS score (i.e. ‘recovered’) at six months follow up if they had been randomised to the exercise intervention. These results emerged despite the fact that substantially more (between two-three fold higher) usual care participants were receiving psychological support at baseline and six month follow up than the exercise group. The benefits of the intervention were seen only at six months follow up but given the longer term adverse effects of postnatal depression for the mother and the development of the baby this is the most critical time to offer treatment.

The magnitude of the difference in mean EPDS scores between the groups was of a moderate size (Affonso, 2000; Matthey, 2004). This is smaller than the overall size found in our meta analysis (Daley et al., 2009) of exercise for postnatal depression (-4.00 EPDS points, 95 % CI: -7.64 to -0.35), but which largely included small studies with high drop-out and less stringent inclusion criteria than used here. In particular, none of the previous trials had recruited women with a *clinical diagnosis* of depression after giving birth. These findings should however be considered in light of the relatively low response (11%) to the initial invitation letter meaning it is possible that women who were more motivated to be active were recruited to the trial.

Our findings are broadly comparable with those reported for other types of treatments for postnatal depression. In the RESPOND trial (n=254) (Sharp et al 2010), which initially randomised women to antidepressants or non-directive counselling but which also allowed women to stop one treatment and start the other or add the other treatment to the first one, reported the proportion of women improving (scoring 13 or less on the EPDS) was 62% for antidepressants and 51% for listening visits 18 weeks after randomisation. These proportions are similar to the proportions considered as improved (56%) and recovered (47%) in our study at six months follow up. The PONDER trial (Morrell et al., 2009) compared the effectiveness of health visitors providing psychological support for one hour per week over eight weeks compared to usual care in women (n=418) considered at risk of postnatal depression (EPDS ≥12). The difference in EPDS score between the groups in was 2.1 points (95% CI -3.3 to -0.9) (adjusted) favouring the intervention group, very similar to our reported mean group difference and confidence intervals.

The exercise intervention group also had higher social support scores at six months indicating that they felt more supported than usual care. Exercise can provide a reason for engaging with others and as part of this intervention participants were encouraged to find

social support, to ask friends/family to support them with their exercise. Findings suggest the intervention group were able to achieve this and plausibly benefitted from doing so. This is important because women with postnatal depression can become isolated at a time when they need support the most. Exercise may be a vehicle by which these women (and health professionals responsible for their care) can engage others in their social networks to offer support. Moreover, most forms of exercise typically involve some level of engagement or connection with others, whether that is by attending exercise groups/classes or taking a walk in the local community for example.

No significant differences between the groups at follow up were recorded for self-reported physical activity; there are several plausible explanations for this. Usual care may have over reported their exercise which is common in the general population. The intervention group however were taught how to accurately assess and report their exercise as a way of self-monitoring their progress so were made more aware of what constitutes different intensities of exercise, potentially resulting in this group reporting their exercise more accurately than usual care. As individuals can increase their exercise without professional support, usual care may have increased their exercise following informed consent. It is also possible that the intervention group did not achieve higher exercise levels compared to usual care.

We are not able to corroborate the self-reported exercise data using data from the Actiheart (objective assessment) because so few women were prepared to wear the device at follow up. Many participants found the Actiheart uncomfortable and/or it interfered with breastfeeding so removed it. However, the logs completed by the exercise group throughout the intervention showed that this group reported increasing their exercise levels over time and maintained participation in at least 150 minutes of moderate/vigorous intensity exercise per week. A real increase in exercise may explain the significant improvements in psychological

outcomes for the women in the intervention group relative to usual care, but other mechanisms are also possible. It is possible the improvements in psychological outcomes are due to non-specific effects associated with the contacts with the PAF, which may have been perceived as therapeutic by women. It may also have been that the PAF acted as a mechanism for women to seek out additional social support to help them with their exercise endeavours, and it is this social support that facilitated improvement in their mental health.

Strengths and limitations of the study

The response rate was relatively low and whilst we did not meet our recruitment target the size of difference seen in EPDS score exceeded the pre-specified effect size in our sample size calculation. Several alternatives strategies were used during the trial to try and improve recruitment but with varied success. To facilitate recruitment future trials will need to involve several centres and/or recruit over several years. A very recent trial of depression (Krusche et al., 2014**)** found that an intensive advertisement campaign that targeted community settings (e.g. adverts on buses, social media and regular radio adverts) was very cost effective. Our trial had minimal involvement from mass media and in future trials this might be a useful supplement to the recruitment strategies used here. This is the largest exercise RCT of any postnatal women considered to be depressed and the only study to recruit women with a clinical diagnosis of depression. This study was extremely rigorous in determining a diagnosis of depression, using a two stage process that resulted in an ICD-10 diagnosis of a major depressive episode or mixed anxiety and depression, rather than only a high score on a screening questionnaire as used in all previous studies. The effect of the intervention on EPDS score are in line with larger high quality studies that have evaluated other treatments for postnatal depression, providing reassurance about the robustness of our findings.

At baseline most women were experiencing a severe/moderate depressive episode and 34% reported thoughts of self-harming demonstrating recruitment of those women clearly in

need of intervention. A large proportion of participants were of non-white ethnicity (37%) and/or living in the two highest IMD deprivation quartiles (79%); typically these populations are very difficult to recruit to mental health treatment trials. Loss to follow up was very low (6% at 6 months, 16% at 12 months). We are not able to be certain of the exact mechanism of effect by which the intervention improved outcomes but we do know the intervention group achieved the behavioural goals asked of them during the exercise intervention. Nevertheless, this trial provides the first steps of evidence and the foundation for further studies on this question in women with clinically defined postnatal depression. Studies that focus on investigating potential mechanisms of effect in this population are needed.

Conclusions

This trial contributes new evidence to indicate that a facilitated exercise intervention that involved encouragement to exercise and to seek out social support to exercise may be an effective treatment for women experiencing postnatal depression, including those at risk of self-harming.

Funding
This study was funded by the National Institute for Health Research School for Primary Care Research. AD is supported by a National Institute for Health Research Senior Research Fellowship. CM and KJ are part funded by the NIHR through the Collaborations for Leadership in Applied Health Research and Care for West Midlands (CLAHRC-WM) programme. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

Acknowledgements
We would like to thank patients and staff who participated in the study and members of the trial steering committee, particularly Professor Nanette Mutrie. We would also like to thank the Mental Health Research Network for their help with recruiting participants and for developing the trial website. We would also like to thanks the Primary Care Clinical Research and Trial Unit at the University of Birmingham for their help with developing the study database and recruiting participants.

Contributors:
AD, KJ, CM, AR, DS, KT, MM and IJ were responsible for the initial proposal, secured funding for the study and drafted the original protocol. RB and SC also contributed to drafting the published protocol. AD as the principal investigator had overall management responsibility for the study. RB as trial coordinator was responsible for the day to day conduct of the study. SC delivered the intervention. RB prepared the data from the actihearts for analysis. AR conducted the sample size calculations, wrote the analysis plan and conducted the statistical analyses. AD wrote the initial draft of the manuscript with assistance from CM and KJ. All authors contributed to and approved the final manuscript. Amanda Daley is the guarantor.

Ethics approval:
Favourable ethical opinion was granted for this study by the Birmingham, East, North and Solihull Research Ethics Committee. Reference 09/H1206/94.

Data sharing
Additional data can be obtained from the corresponding author for the purposes of secondary research.

Conflict of interest disclosures
None.

ISRCTN trial registration: CCT-NAPN-13286.

438 **References**

439

440 **Affonso DD, De AK, Horowitz JA, Malberry LJ**. (2000). An international study exploring
441 levels of postpartum depressive symptomatology. *Journal of Psychosomatic Research* **49**,
442 207– 216.

443

444 **American College of Obstetrics and Gynecology (ACOG) (**January 2002) Committee
445 Obstetric Practice. ACOG Committee opinion. Number 267, Exercise during pregnancy and
446 the postpartum period. *Obstetrics and Gynecology* **99**, 171-173.

447

448 **American Psychiatric Association**. (2000). *Diagnostic and Statistical Manual of Mental*
449 *Disorders*. American Psychiatric Association: Washington DC.

450

451 **Armstrong K, Edwards H**. (2003). The effects of exercise and social support on mothers
452 reporting depressive symptoms: A pilot randomized controlled trial. *International Journal of*
453 *Mental Health Nursing* **12**, 130-138.

454

455 **Beck C**. (1995). The effects of postpartum depression on maternal-infant interactions: A
456 meta analysis. *Nursing Research* **44**, 298-304.

457

458 **Blamey RV, Daley AJ, Jolly K. (2012).** Exercise for postnatal psychological outcomes: a
459 systematic review and meta-analysis. *Lancet* **38,** S25.

460

461 **Bostic TJ, Rubio DM, Hood M**. (2000). A validation of the subjective vitality scale using
462 structural equation modeling. *Social Indicators Research* **52**, 313-324.

463

464 **Brooks R**. (1996). EuroQol: the current state of play. *Health Policy* **37**, 53-72.

465

466 **Chief Medical Officers of England, Scotland, Wales, and Northern Ireland.** (2011). Start
467 Active, Stay Active. A report on physical activity for health from the four home countries'
468 Chief Medical Officers. Department of Health.
469 (http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuida
470 nce/DH_128209). Accessed 09 June 2014.

471

Cooper PJ, Murray L. (1995). Course and recurrence of postnatal depression. Evidence for the specificity of the diagnostic concept. *British Journal of Psychiatry* **166**, 191–195.

Cox J, Holden J, Sagovsky R. (1987). Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry* **150**, 782-786.

Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, Pratt M, Ekelund, U, Yngve, A, Sallis, JF, Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise* **35**, 1381-1395.

Daley AJ, Jolly K, MacArthur C. (2009) The effectiveness of exercise in the management of postnatal depression: Systematic review and meta-analysis. *Family Practice* **26**,154-162.

Daley AJ, Jolly K, Sharp DJ, Turner KM, Blamey RV, Coleman SJ, McGuinness M, Roalfe AK, Jones I, MacArthur C. (2012). The effectiveness of exercise as a treatment for postnatal depression: Study protocol. *Biomed Central Pregnancy and Childbirth* **12**, 45.

Daley AJ, MacArthur, C, Winter H. (2007). The role of exercise as a treatment for postnatal depression: A review. *Journal of Midwifery and Women's Health* **52**, 56-62.

Daley AJ, Winter HR, Grimmett, C, **McGuinness M, McManus R, MacArthur C.** (2008). The feasibility of an exercise intervention in depressed postnatal women: A pilot randomised controlled trial. *British Journal of General Practice* **58**, 178-183.

Dennis C-L, Hodnett E. (2007). Psychosocial and psychological interventions for treating postpartum depression. *Cochrane Database of Systematic Reviews* **4**, Art.No: CD006116.

Gaynes BN, Gavin N, Melzer-Brody S, Lohr KN, Swinson T, Gartlehner G, Brody, S. (February 2005). Perinatal depression: prevalence, screening accuracy and screening outcomes. Summary, Evidence Report/Technology Assessment no.119. Prepared by RTI-University of North Carolina Evidence based Practice Center under contract No.290-02-

0016.). AHRQ Publication No.05-E006-1. Rockville, MD: Agency for Healthcare Research and Quality.

Heh S, Huang L, Ho S, Fu Y, Wang L (2008). Effectiveness of an exercise support program in reducing the severity of postnatal depression in Taiwanese women. *Birth* **35**: 60-65.

International Classification of Disease and Health Related Problems version 10. World Health Organisation, (2011). (http://www.who.int/classifications/icd/ICD10Volume2_en_2010.pdf.) Accessed 09 June 2009.

Jacobson NS, Traux P (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting Clinical Psychology **59**, 12-19.

Jenkins R, Lewis G, Bebbington P, Brugha T, Farrell M, Gill B, Meltzer, H. (1997). The National Psychiatric Morbidity Surveys of Great Britain – initial findings from the Household Survey. *Psychological Medicine* **27**, 775–89.

Kumar R, Robson KM, Smith AM. (1984). Development of a self-administered questionnaire to measure maternal adjustment and maternal attitudes during pregnancy and after delivery. *Journal of Psychosomatic Research* **28**, 43-51.

Krusche A, Rudolf von Rohr I, Muse K, Duggan D, Crane C, Williams JM. An evaluation of the effectiveness of recruitment methods: the staying well after depression randomized controlled trial. *Clin Trials* 2014; **11**: 141-9.

Lewis G. (1994). Assessing psychiatric disorder with a human interviewer or a computer. *Journal of Epidemiology and Community Health* **48**, 207–10.

Lewis BA, Gjerdingen DK, Avery MD, Sirard JR, Guo H, Schuver, K., Marcus, B.H. (2014) Randomized trial examining a physical activity intervention for the prevention of postpartum depression: The healthy mom trial. *Mental Health and Physical Activity* **7,** 42-49.

Lewis G, Pelosi A, Araya R, Dunn G. (1992). Measuring psychiatric disorder in the community: a standardised assessment for use by lay interviewers. *Psychological Medicine* **22**, 465-488.

Matthey S. (2004). Calculating clinically significant change in postnatal depression studies using the Edinburgh Postnatal Depression Scale. *Journal of Affective Disorders* **78**, 269–72.

McLennan D, Barnes H, Noble M, Davies J, Garratt E. (March 2011). *The English Indices of Deprivation 2010*. Department for Communities and Local Government. London. (https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6320/1870718.pdf). Accessed 09 June 2010.

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

Murray L. (1992). The impact of postnatal depression on infant development. *Journal of Child Psychology and Psychiatry* **33**, 543-561.

National Institute for Health and Clinical Excellence. (2007). *CG45 Antenatal and postnatal mental health. Clinical Management and Service Guidance*. London: NHS

O'Hara M, Swain A. (1996). Rates and risk of postpartum depression–a meta analysis. *International Review of Psychiatry* **8**, 37-54.

Sallis JF, Grossman RM, Pinski RB, Patterson TL, Nader PR. (1987). The development of scales to measure social support for diet and exercise behaviours. *Preventive Medicine* **16**, 825-836.

Sallis JF, Pinski RB, Grossman RM. (1988). The development of self-efficacy scales for health-related diet and exercise behaviours. *Health Education Research* **3**, 283-292.

Sharp DJ, Chew-Graham C, Tylee A, Lewis G, Howard L, Anderson I, Abel K, Turner KM, Hollinghurst SP, Tallon D, McCarthy A, Peters TJ. (2010). A pragmatic randomised controlled trial to compare antidepressants with a community-based psychosocial intervention for the treatment of women with postnatal depression: the RESPOND trial. *Health Technology Assessment* **14** (43).

Singleton N, Lewis G. (2003). *Better or worse: A longitudinal study of the mental health of adults living in private households in Great Britain*. Office for National Statistics, Her Majesty's Stationery Office, London.

Turner K, Chew-Graham C, Folkes L, Sharp D. (2010). Women's experiences of health visitor delivered listening visits as a treatment for postnatal depression: a qualitative study. *Patient Education and Counseling* **78**, 234-9.

Turner K, Sharp D, Folkes L, Chew-Graham C. (2008). Women's views and experiences of antidepressants as a treatment for postnatal depression: a qualitative study. *Family Practice* **25**, 450-5.

Ware JE, Kosinski M, Keller SD. (1996). A 12-Item Short-Form Health Survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care* **34**, 220-233.

595 **Table 1: Baseline characteristics of randomised participants**

	Exercise (N=47)	**Usual care (N=47)**
	n (%)	n (%)
Age (years) mean (sd)	31.7 (5.3)	29.3 (5.7)
EPDS [0-30] mean (sd)	17.3 (3.0)	17.5 (3.7)
EPDS score* 13-16	20 (43)	20 (43)
17+	27 (57)	27 (57)
Thoughts of self-harming	14 (30%)	18 (38%)
IMD quartile 1 (least deprived)	3 (6)	4 (9)
2	7 (15)	6 (13)
3	13 (28)	9 (19)
4 (most deprived)	24 (51)	28 (60)
Ethnic origin		
White	27 (57)	32 (68)
Mixed	3 (6)	2 (4)
Indian	3 (6)	3 (6)
Pakistani	7 (15)	5 (11)
Bangladeshi	1 (2)	0 (0)
Black-African	2 (4)	1 (2)
Black-Caribbean	0 (0)	1 (2)
Chinese	1 (2)	0 (0)
Other	3 (6)	3 (6)
Number of children mean (sd)	2.2 (1.2)	2.2 (1.4)
Number with children aged 2-5 years	25 (53)	22 (47)
Number with children aged over 5 yrs	14 (30)	12 (26)
Other occupants - Husband	40/45 (89)	41/47 (87)
Relative	7/30 (23)	6/26 (23)
Step-children	1/25 (4)	2/23 (9)
Age of baby at randomisation (days) mean (sd)	117.3 (26.5)	121.8 (27.9)
Height (m) mean (sd)	1.63 (0.06)	1.64 (0.06)
Weight (kg) mean (sd)	76.2 (13.4)	76.8 (17.0)
BMI mean (sd)	28.8 (5.0)	28.5 (5.8)
Smoker	9/46 (20)	9/47 (19)
Experiencing long term illness	7/47 (15)	7/47 (15)
Prescribed antidepressants before or during pregnancy	18/39 (46)	12/37 (32)
Currently taking antidepressants	10/46 (22)	10/47 (21)
Currently having counselling/psychological support	3/46 (7)	10/47 (21)
Currently taking antidepressants plus having counselling/psychological support	1(2)	6(13)
Employment status:		
Paid	17/46 (37)	25/47 (53)
Self employed	0/46 (0)	2/47 (4)
Unemployed	7/46 (15)	5/47 (11)
Student	1/46 (2)	2/47 (4)
Looking after home/family	20/46 (43)	13/47 (28)

Sick/disabled	0/46 (0)	0/47 (0)
Other	1/46 (2)	0/47 (0)
Years living in area median (IQR)	5 (2 to 10.5)	4 (2 to 10)
Have practical help at home	35/47 (74)	36/46 (78)
Have contact with local people with a baby	22/47 (47)	27/44 (61)
Have someone to talk to	39/47 (83)	35/46 (76)
Have help from husband [0-10] mean (sd)	5.4 (3.3)	5.7 (2.7)
Currently breast feeding	16/47 (34)	20/46 (44)
Type of delivery		
Normal vaginal	27/47 (57)	31/46 (67)
Instrumental vaginal	6/47 (13)	6/46 (13)
Elective caesarean	7/47 (15)	5/46 (11)
Emergency caesarean	7/47 (15)	4/46 (9)
Exercise per wk pre-pregnancy (hours) median (IQR)	0 (0 to 5)	0 (0 to 4)

*Stratification variable. Higher score on EPDS indicates high levels of probable depression

598 **Table 2: Baselines questionnaires and measurements**

Outcomes		**Intervention** Mean (sd)	N	**Usual care** Mean (sd)	N
Questionnaires					
Vitality scale [1-7]		2.8 (0.7)	46	2.8 (0.6)	47
PCS-12 [0-100]	mean (sd)	52.7 (7.9)	47	51.0 (9.4)	47
MCS-12 [0-100]	mean (sd)	30.8 (7.9)	47	31.2 (7.9)	47
EQ-5D [-0.59-1.0]	mean (sd)	0.68 (0.19)	47	0.68 (0.22)	46
	median (IQR)	0.73 (0.62 to 0.81)		0.69 [0.69 to 0.85]	
MAMA (body image) [10-40]		22.5 (4.4)	44	22.0 (3.9)	41
Social support [8-24]		20.1 (4.0)	47	19.5 (4.4)	47
Social support and exercise					
Family participation [10-50]	mean (sd)	15.1 (6.5)	44	14.9 (5.4)	44
	median (IQR)	12 [10 to 20]		12.5 [10 to 20.5]	
Family rewards [3-15]	mean(sd)	3.4 (1.2)	45	3.2 (0.8)	45
	median (IQR)	3 [3 to 3]		3 [3 to 3]	
Friends participation [10-50]	mean (sd)	13.0 (5.6)	44	12.8(4.3)	44
	median (IQR)	10.5 [10 to 14]		10 [10 to 14]	
Exercise confidence					
Sticking to it score [8-40]	mean(sd)	20.9 (5.9)	37	19.7 (4.8)	40
Making time for it score [4-20]	mean(sd)	13.0 (3.4)	35	11.8 (2.8)	36
IPAQ					
Vigorous (MET-mins/wk)	mean(sd)	264.7 (762.9)	47	163.6 (546.7)	44
	median (IQR)	0 [0 to 0]		0 [0 to 0]	
Moderate (MET-mins/wk)	mean(sd)	453.9 (927.9)	46	59.1 (167.1)	46
	median (IQR)	0 [0 to 360]		0 [0 to 0]	
Walking (MET-mins/wk)	mean(sd)	950.5 (1089.6)	43	895.1 (1106.6)	44
	median (IQR)	495 [198 to 1386]		585.8 [99 to 1254]	
Total (MET-mins/wk)	mean(sd)	1725.4 (1864.7)	43	1122.1 (1242.9)	43
	median (IQR)	918 (396 to 3108)		594 (99 to 1254)	
Sitting (hrs/day)	mean(sd)	5 (4.1)	32	5.3 (4.9)	28
	median (IQR)	4 [2.5 to 6]		4 [3 to 5]	

599 Higher scores on all questionnaire based outcomes except EPDS and sitting time indicate
600 more positive health/behaviours
601
602

Table 3: Six and 12 month post randomisation follow up data for outcomes

	Follow-up (months) post randomisation	Randomisation Group							
		Exercise		Usual care					
		mean (sd)	n	mean (sd)	n	Difference in adjusted* means (95% CI)	p value	Difference in adjusted** means (95% CI)	p value
Primary outcome:									
EPDS [0-30]	6	12.51 (5.46)	43	14.67 (4.86)	42	-2.04 (-4.11 to 0.03)	0.053	-2.26 (-4.36 to -0.16)	0.035
	12	12.02 (5.29)	41	12.55 (5.17)	38	-0.95 (-3.16 to 1.25)	0.40	-1.39 (-3.69 to 0.92)	0.24
Weight (kg)	6	75.06 (13.20)	42	77.36 (17.36)	42	0.38 (-1.82 to 2.58)	0.74	0.79 (-1.41 to 3.00)	0.48
	12	75.40 (16.05)	41	77.11 (18.39)	38	1.39 (-2.12 to 4.91)	0.44	1.36 (-2.36 to 5.09)	0.47
BMI (kg/m^2)	6	28.61 (5.38)	42	28.61 (5.98)	41	0.17 (-0.67 to 1.01)	0.69	0.22 (-0.61 to 1.06)	0.45
	12	28.75 (6.99)	41	28.77 (6.29)	37	0.53 (-0.90 to 1.96)	0.47	0.37 (-1.15 to 1.88)	0.63
PCS-12 [0-100]	6	51.34 (9.02)	42	51.59 (8.48)	42	-1.10 (-4.32 to 2.11)	0.50	-0.06 (-2.95 to 2.84)	0.97
	12	52.16 (9.16)	40	51.6 (8.57)	38	-0.49 (-3.76 to 2.78)	0.77	0.10 (-3.17 to 3.37)	0.95
MCS-12 [0-100]	6	41.45 (9.99)	42	37.90 (10.30)	42	3.38 (-0.74 to 7.51)	0.11	3.45 (-0.78 to 7.69)	0.11
	12	41.60 (12.13)	41	41.02 (12.36)	38	1.16 (-3.65 to 5.96)	0.64	1.97 (-3.13 to 7.07)	0.45
EQ-5D # [-0.59-1.0]	6	0.78 (0.21)	41	0.72 (0.22)	41	0.07 (-0.02 to 0.15)	0.12	0.07 (-0.02 to 0.15)	0.11
	12	0.81 (0.21)	40	0.78 (0.23)	38	0.05 (-0.03 to 0.14)	0.22	0.06 (-0.03 to 0.15)	0.22
MAMA (Body image) [10-40]	6	23.73 (5.20)	40	22.53 (4.10)	40	0.53 (-1.01 to 2.07)	0.50	0.60 (-0.93 to 2.14)	0.44

	Follow-up (months) post randomisation	Randomisation Group							
		Exercise		Usual care					
		mean (sd)	n	mean (sd)	n	Difference in adjusted* means (95% CI)	p value	Difference in adjusted** means (95% CI)	p value
	12	24.43 (5.18)	37	23.94 (4.83)	36	-0.11 (-1.99 to 1.76)	0.91	-0.03 (-2.04 to 1.97)	0.98
Vitality [1-7]	6	3.53 (0.95)	39	3.18 (0.87)	40	0.36 (-0.04 to 0.76)	0.08	0.41 (0.007 to 0.82)	0.054
	12	3.70 (1.11)	41	3.48 (1.14)	38	0.31 (-0.14 to 0.76)	0.18	0.42 (-0.06 to 0.90)	0.09
Social support [8-24]	6	20.80 (3.59)	41	18.93 (4.93)	41	2.10 (0.78 to 3.43)	0.002	2.24 (0.93 to 3.54)	0.001
	12	20.80 (3.68)	40	19.73 (5.14)	37	1.25 (-0.20 to 2.71)	0.09	1.38 (-0.15 to 2.90)	0.08
Social support for exercise Family participation [10-50]	6	16.87 (7.44)	38	15.48 (6.54)	40	2.23 (-0.94 to 5.40)	0.17	2.00 (-1.49 to 5.48)	0.26
	12	16.88 (7.87)	41	18.67 (9.53)	36	-0.95 (-4.28 to 2.38)	0.58	-0.88 (-4.30 to 2.54)	0.62
Family rewards and punishment [3-15]	6	3.60 (1.68)	40	3.55 (1.30)	40	0.11 (-0.41 to -0.64)	0.68	0.16 (-0.37 to 0.69)	0.55
	12	3.46 (1.47)	41	4.03 (1.65)	38	-0.54 (-1.06 to -0.02)	0.04	-0.42 (-0.96 to 0.11)	0.12
Friend participation [10-50]	6	16.57 (9.08)	37	14.75 (8.11)	40	1.62 (-2.12 to 5.37)	0.40	2.18 (-1.85 to 6.20)	0.29
	12	14.55 (7.13)	38	16.73 (8.36)	37	-2.83 (-5.65 to -0.005)	0.050	-2.30 (-5.15 to 0.55)	0.11

	Follow-up (months) post randomisation	Randomisation Group							
		Exercise		Usual care					
		mean (sd)	n	mean (sd)	n	Difference in adjusted* means (95% CI)	p value	Difference in adjusted** means (95% CI)	p value
Exercise confidence ECS sticking to it [8-40]	6	22.09 (7.93)	33	20.22 (6.89)	32	2.41 (-1.71 to 6.52)	0.25	3.27 (-1.13 to 7.66)	0.15
	12	23.29 (8.54)	31	22.03 (8.24)	30	0.88 (-3.30 to 5.07)	0.68	2.03 (-2.52 to 6.57)	0.38
ECS making time for it [4-20]	6	12.21 (4.26)	34	11.07 (3.38)	30	0.27 (-1.89 to 2.44)	0.81	0.91 (-1.25 to 3.08)	0.41
	12	12.81 (3.79)	31	12.15 (4.35)	26	-0.73 (-1.43 to 2.90)	0.51	1.32 (-1.03 to 3.68)	0.27

* adjusted by baseline value and baseline EPDS

** adjusted by baseline value, baseline EPDS, age, weight and ethnicity

bootstrapped confidence intervals and p values

Higher scores on all questionnaire based outcomes except EPDS and sitting time indicate more positive health/behaviours

Table 4: Proportions improved and recovered on the EPDS six and 12 months post randomisation

	Randomisation Group			
Outcome	**Exercise**	**Usual care**	**Difference (95% CI)**	p value
Improved at 6 months (EPDS score less than 13)	24/43 (55.8%)	16/42 (38.1%)	17.7% (-3.2% to 38.6%)	0.10
Improved at 12 months (EPDS score less than 13)	23/41 (56.1%)	17/38 (44.7%)	11.4% (-10.6% to 33.3%)	0.31
Recovered at 6 months (EPDS dropped by 4 points and score less than 13)*	20/43 (46.5%)	10/42 (23.8%)	22.7% (3% to 42.4%)	0.03
Recovered at 12 months (EPDS dropped by 4 points and score less than 13)*	21/41 (51.2%)	14/38 (36.8%)	14.4% (-7.3% to 36.0%)	0.20

*Criteria for determining clinically important change (Jacobson and Truax 1991; Matthey 2004).