Reviewer's report

Title: Preferred intensity exercise for adolescents receiving treatment for depression: A pragmatic randomised controlled trial

Version: 1 Date: 16 June 2015

Reviewer: Nicola Wiles

Reviewer's report:

The authors report the findings from a randomised controlled trial of an exercise intervention for adolescents with depression who are in contact with mental health services. This is an important area and the current literature is contradictory and previous trials are small and have methodological weaknesses. I have a number of key issues that warrant further attention.

Major compulsory revisions

1. The trials register entry https://clinicaltrials.gov/ct2/show/NCT01474837 suggests that the trial intervention was exercise plus motivational interviewing. There is no mention of the latter in the description of the intervention. Although the authors do state that two project staff members “exercised and interacted with participants in all sessions”. This needs clarification.

2. In the description of the analysis, the authors suggest that the primary model for the primary outcome included an interaction between treatment by time. However in the reporting of the results, the authors only present the findings for outcomes post-intervention and at 6 months. This is inconsistent and requires clarification. Was there any evidence of a differing effect of time over time? Given the numbers, I suspect that any such formal test of interaction would be substantially underpowered.

3. In the results section, the authors talk in terms of findings that were or were not ‘statistically significant’. As the authors themselves acknowledge, they did not recruit the number as specified in the original sample size calculation in order to be adequately powered to detect the pre-specified difference. It is more helpful, in general, to talk in terms of the strength of evidence based on p values and use confidence intervals to describe whether the results exclude a meaningful difference between groups (see Sterne, BMJ 2001;322:226). This needs to be addressed through this section.

4. In the discussion, the authors suggest that a lack of power explains the null finding post-intervention (6 weeks) and that the treatment effect at 6 months is a novel finding. However, there is substantial and, importantly, differential attrition. 43% of those in the intervention group and 58% of those in the TAU group are lost to follow-up at 6 months. Therefore I have concerns about the potential for bias in the results presented. The authors conduct a number of comparisons to test for differences in drop-out between groups and in terms of baseline
characteristics that predict missingness, and finding no significant differences conclude that there is no bias as the missing mechanism is MAR. [The latter is incorrect, if there are truly no factors that predict the missingness then the mechanism is MCAR rather than MAR.] However, given the small sample, basing the conclusions regarding patterns of missingness on statistical significance may be misleading. Given the substantial amount of missing data, particularly at the 6 month point, the authors need to conduct additional sensitivity analyses to determine the potential for bias. Without doing so, I believe that there is a substantial risk that the authors may be drawing a conclusion regarding a long-term effect of the intervention that is not justified.

Minor essential revisions

5. There is no mention of the CONSORT guidelines in the reporting of this trial and it would be helpful to ensure that the manuscript fully aligns with the required reporting standards for RCTs.

6. The sample size calculation indicates that the target sample size was 158. However, only 87 individuals were randomised. There is no mention of the challenges that recruitment posed in this trial and brief details would be helpful.

7. In discussing the literature in adults, the authors should reference the latest version of the Cochrane review of exercise for depression http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD004366.pub6/abstract;jsessionid=2852DF3AF04907E0FEBA6AA1ECF44DED.f02t02

8. I struggled with the term ‘preferred intensity’. I think it needs to be clearer how this was agreed. There is no indication of how long each exercise session was. This needs adding.

9. The authors state that intervention “engagement” was measured using a scale that measured exertion and heart rate readings. I would consider “engagement” to be attendance at sessions and better quantified as the number of sessions attended. Exertion and heart rate are linked with how hard the person is exercising and the level of intensity of exercise. The authors also describe how the participants were asked to report exertion at intervals throughout the exercise session but it is not clear on how many occasions this happened and whether this was consistent across sessions.

10. Under eligibility criteria, 3rd line, typo ‘Imventory’.

11. Under ‘randomisation and allocation concealment’, further information is required as to who gained informed consent for trial participation – was this the referring clinician or a member of the research team?

12. In the description of the LTEQ, it would be helpful to have a description of what the categories ‘active’, ‘moderately active’ and ‘insufficiently active’ equate to in relation to national guidelines of PA levels for children.

13. The total number of participants assessed for eligibility is given as 128 – is this the same as the number of referrals that was received by the team? I
wondered whether there were any individuals who were referred but who did not agree to be screened for eligibility.

14. The description of the average number of sessions attended should include those who did not attend any sessions.

15. The presentation of data in the Table 3 requires attention. The means for the intervention and TAU groups are presented for the various outcomes. The ‘difference in reduction’ is incorrect. This is a (between-group) difference in mean scores. For the EQ-5D-5L, the authors present the median (IQR) values but then present what appears to be a difference in means. This is inconsistent and needs to be addressed. There should only be one p values for the comparison of LTEQ scores between the groups (rather than for each level).

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests