Reviewer’s report

**Title:** Effects of an 18-week exercise programme started early during breast cancer treatment: a randomised controlled trial in daily clinical practice

**Version:** 1  
**Date:** 16 February 2015  
**Reviewer:** Karen Steindorf

**Reviewer’s report:**

This, in general, well-written manuscript describes a randomized two arm trial in which patients in the early phase of treatment for breast cancer received usual care or a mixed aerobic and muscle strength training interventions. The findings from this well-conducted study are important, particularly with regard to minimization of physical fatigue through exercise therapy.

1) Major Compulsory Revisions  
The author must respond to these before a decision on publication can be reached. For example, additional necessary experiments or controls, statistical mistakes, errors in interpretation.

**Title, Background, Methods, and Discussion:**

It was not clear enough to the reviewer where the relevant point for the extension “in daily clinical practice” in the title is. Also the point with “Explanatory trials” in the background section is not taken well enough. What are the major differences to other trials? At least most of the cited publications were also embedded in normal clinical routines. Did that impact the use of standard procedures for this study? If so, please discuss possible effects on the internal and external study validity.

**Results and Discussion:**

As the PACT study has been a multi-center trial, some discussion of this issue would be appreciated. Did this approach lead to more heterogeneity compared to other studies, and if so, with regard to which factors. Does this result in a higher generalizability of results?

**Description of the study population in the methods and results sections**

Please provide more details on the planned study population. All information should be given at one place in the method section. Exercise started within 6 weeks after diagnosis (Abstract) and patients needed to be “scheduled for chemotherapy”. These are very broad inclusion criteria. Some treatment schemes have a scheduled chemotherapy but it would not happen within the intervention window of this study. Were these patients included? Were there any patients that were not under CTx during the intervention period?

Also, the description of the final study population in the results section should be
improved. Only very late in the discussion the reader learns that patients were heterogeneous with regard to treatment status at the time of the baseline assessment. Patients were equally distributed over 3 groups at baseline, patients who started CTx, patients who started RTx, and patients who had not started adjuvant therapy. Please include this information into Table 1. For how long are patients treated with RTx before CTx in general? There seems to be some variation in treatment schemes between countries.

The reviewer has some concern with these three groups and feels that this point could be elaborated in more detail. As we know from many studies, CTx has the strongest impact on fatigue, thus the three patient groups are presumably very different with regard to fatigue at baseline. As stratification factor only “radiotherapy yes/no before CTx” was used. However, this means that women with or without CTx were pooled and separated from RTx patients. What was the reasoning for this stratification? The expectation from a clinical perspective would be that ongoing CTx or baseline fatigue would have been better stratification factors, especially as fatigue was the primary outcome.

The authors somehow addressed the issue by performing a sensitivity analyses on CTx before randomization and reported non-significant interactions. But this result might just be explained by a restricted power to test this interaction. This general issue should be discussed as limitation of the study, not only as a strength.

Please also provide more details on the surgical status of the patients. In setting, 1st paragraph, patients with immediate reconstruction are mentioned. What types of surgery are summarized under this term? Do you cover breast-conserving surgeries with this term?

How many patients had neoadjuvant CTx and how many adjuvant CTx? Do you see an issue for interpreting your results?

Please also provide some information on several time windows, for example, time between diagnosis and randomization, time between start of first treatment and randomization etc. Were patients with RTx before CTx included later than the CTx only patients?

Control group

The reviewer wonders how surprising the contamination of the control was. In the section “Intervention” it reads that exercise programmes are offered routinely to cancer patients having completed primary treatment. What is the situation in The Netherlands for patients during treatment?

Abstract (Results) and Statistical Methods:

Fatigue was defined as primary endpoint? Does this refer to general fatigue or to any subscale? Was the analysis at 18-weeks or at 36-weeks planned as primary analyses of the trial or what was the primary hypothesis of the trial?

Abstract, Results, 1st sentence:
Please restructure this sentence to improve readability and correctness. Maybe splitting into two parts and using more different separators within the bracket would help. Please be more specific on the fact that the result for general fatigue was not significant.

Abstract
The presentation of the results on the patients' fitness is misleading. It is not clear that this statement summarizes results of several measured parameters. The effect sizes do not range from 0.25-0.45 but from 0.02-0.45 with 6 out of 16 parameters reaching statistical significance (see Table 4).

Abstract (Conclusions):
The last sentence needs rephrasing. Shortening of the first part but being more precise in the second might help.

Background:
Please check if reference 1 is valid for Western populations and provide a more recent source for Reference 3. Furthermore, the meta-analysis of Brown-F on exercise and cancer-related fatigue should be added as important source.
The placing and the corresponding text on references 9 and 10 raises questions as before the quality of previous studies as well as lacking ITT analyses are mentioned. Please be more precise if these criteria are critical for these two publications too.
Last sentence: either delete or remove to the methods or discussion section.

Intervention:
What was the proportion of time within one training session with regard to the 60 minutes for aerobic and strength training? How standardized was the setting for the aerobic training? What type of exercise was it?

Sample size:
Why were more patients recruited than planned?
Please be more precise for the term “tumor receptor” … and add “status”

Statistical methods:
In general mixed linear models it is possible to include all collected data even if patients were non-completers. But if non-completers differ systematically from completers results may be biased. As it is reported in the results section that non-completers had higher fatigue, it would be interesting if results changed if standard techniques for handling of missing data were applied (complete-case, multiple imputation)?

What happens if you adjust for a variable that mirrors the three pre-treatment groups? Is there a difference between adjustment and testing interaction effects in your model?

Results
Participants + Figure 1:
• What does “completed the trial” exactly means.
• 2nd paragraph, 1st line: Not 13 but 25 participants from the control group were lost according to Figure 1. Were the proportions of missing data different for the different variables?

Fatigue:
Please rephrase the sentence starting in line 5 and do not use “all” participants if you refer to mean values for both groups.
Is the second last sentence in agreement with table 2?

Tables 2-4 and S1:
• The column headings “18 weeks post intervention” are misleading. The assessment is 18 weeks after baseline and post intervention but not 18 weeks away from the intervention period.
• In the footnote it reads that the within-group differences in the table are based on 102 participants in each group that had baseline values. However, the within-group differences refer to changes from baseline to the other timepoints where some of these participants did not have data. How could you include all 204 participants?

Table 4:
This table needs major changes to increase readability. “Ext 60d” is too technical and will not help many readers from BMC medicine. Muscle strength should read “Leg muscle strength”.

Adherence to the study protocol:
Second last sentence: This may be an explanation. However, if this is important it needs to be elaborated in more detail along Table S2. Maybe, this issue would be better placed in the discussion.

Discussion:
4th paragraph: The BEATE did not exclude patients with baseline depression in general. There are tables where depressed patients were excluded but these were only secondary analyses.
As some of your participants received not only chemotherapy but also radiotherapy your results should also be set into the context of those studies (e.g. BEST study). As in both settings exercise effects have recently been reported this may help discussing the effect of the three pre-treatment groups.

Discussion, 7th paragraph:
The presented data on the study population with regard to previous treatment needs to be given earlier. The reviewer does not agree with this argumentation. There are also many disadvantages due to this mixture that needs to be discussed (see above).
2) Minor Essential Revisions
The author can be trusted to make these. For example, missing labels on figures, the wrong use of a term, spelling mistakes.

Please delete “also” in the Abstract, Results, 4th line
Please delete “also” in the Background, 2nd last sentence
Intervention, line 6: at least one “,” is missing.

Table 1
Please correct the unit of the BMI.
Please be more congruent with capitalizing letters.

Table 3:
2nd block in the table as a wrong text (18# 36)

3) Discretionary Revisions
These are recommendations for improvement which the author can choose to ignore. For example clarifications, data that would be useful but not essential.

You may want to check on the frequency and necessity of the term “also” throughout the manuscript.

Discussion, 5th paragraph, last line: I suggest to replace “can” by “may”.

Acknowledgements: N travier # capitalize T

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare that I have no competing interests