

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Cost-effectiveness analysis of an 18-week exercise programme for breast and colon cancer patients undergoing adjuvant chemotherapy: the randomized PACT study
AUTHORS	May, Anne; Bosch, Marcel; Velthuis, Miranda; van der Wall, Elsken; Steins Bisschop, Charlotte; Los, Maartje; Erdkamp, Frans; Bloemendal, Haiko; de Roos, Marnix; Verhaar, Marlies; ten Bokkel Huinink, Daan; Peeters, Petra; Wit, Ardine

VERSION 1 - REVIEW

REVIEWER	Line Oldervoll NTNU, Norway
REVIEW RETURNED	08-May-2016

GENERAL COMMENTS	<p>Concerning Manuscript ID: bmopen – 2016- 012187</p> <p>Title: Cost-effectiveness analysis of an 18-week exercise programme for breast and colon cancer patients undergoing adjuvant chemotherapy: the randomized PACT study</p> <p>The aim of this paper is to report data on cost-effectiveness of an 18 week exercise programme for breast and colon cancer during treatment with adjuvant chemotherapy.</p> <p>The paper is well-written, interesting and easy to read. It's an important study and one of few who have addressed cost-effectiveness of physical exercise programmes for cancer patients in addition to effects.</p> <p>The results show that for colon cancer, the cost-effectiveness analysis showed beneficial effects of the exercise intervention. For breast cancer patients the results did not indicate that the exercise intervention was cost-effective.</p> <p>The limitation of the study is the low number of patients included, especially for the colon cancer group. As the authors state in the discussion, this might have caused a false positive result and the study needs to be replicated in a study with a greater number of patients.</p> <p>I believe this paper deserves to be published as it is written.</p>
-------------------------	---

REVIEWER	Patricia Herman RAND Corporation
REVIEW RETURNED	08-Jun-2016

GENERAL COMMENTS	<p>Overall, this is a very well-done CUA. Minor comments:</p> <p>1) Single imputation was used for missing costs and missing EQ-5D scores, yet little information is given on the amount of missing data (the number of imputations needed) and the impact of these imputations--i.e., a full case sensitivity analysis would have been</p>
-------------------------	---

	<p>helpful.</p> <p>2) The manuscript says that "a bootstrap analysis" was used. However, the gold standard is BCa (bias-corrected and accelerated) bootstrap. Was this done here, and if so, please state that, and if not, why not?</p> <p>3) On page 10 in the section labeled QALYs, all results are reported as a change in the number of QALYs. However, Table 3 shows the first two of these values for each type of cancer as "Average EQ-5D" scores, and from the results shown there it looks like that is what they are. The QALY values that are labeled as such in Table 3 are reported correctly in the text.</p> <p>4) On Table 3 both the EQ-5D score and the EQ-5D VAS are reported. I would recommend that the EQ-5D VAS values be removed from the table because they are inappropriate for use in a societal perspective analysis. The EQ-5D VAS measures HRQoL from the patient's perspective, whereas, the EQ-5D has (Danish) societal preference weights applied to give HRQoL for the reported health states from the societal perspective. I would also recommend removal of any mention of the EQ-5D VAS from the text unless the authors want to explain its use.</p> <p>5) Also, taking the EQ-5D VAS numbers out of Table 3 would give more room to report the averages of the full set of 9 every-4-week EQ-5D scores. Given the dramatic differences seen between the BCA and CCA cohort, providing more transparency here would help. I also wonder if the authors applied the full AUC approach as recommended by Manca (Health Econ. 14: 487–496 (2005)) including the adjustment for differences in baseline values. If not, that is recommended.</p> <p>6) On the bottom of page 13 a statement is made that "about 50% of control group patients also reported high levels of physical activity." Given the dramatic differences seen between the QALY results for the BCA and CCA cohorts, it is important to indicate whether both these groups experienced this problem equally.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Comment reviewer 1:

The paper is well-written, interesting and easy to read. It's an important study and one of few who have addressed cost-effectiveness of physical exercise programmes for cancer patients in addition to effects. The results show that for colon cancer, the cost-effectiveness analysis showed beneficial effects of the exercise intervention. For breast cancer patients the results did not indicate that the exercise intervention was cost-effective. The limitation of the study is the low number of patients included, especially for the colon cancer group. As the authors state in the discussion, this might have caused a false positive result and the study needs to be replicated in a study with a greater number of patients. I believe this paper deserves to be published as it is written.

Answer to reviewers 1's comment:

We thank the reviewer for her positive response. We agree that our results need to be replicated and this is stated in the discussion.

Comments reviewer 2

Overall, this is a very well-done CUA.

Authors' answer:

Thank you. We also appreciate the minor (methodological) comments mentioned below, which

improved our manuscript.

Minor comments:

1) Single imputation was used for missing costs and missing EQ-5D scores, yet little information is given on the amount of missing data (the number of imputations needed) and the impact of these imputations--i.e., a full case sensitivity analysis would have been helpful.

Authors' answer:

The amount of missing costs and missing EQ-5D scores was low (<10%) during the 18-week intervention period. During the 18 week follow-up (after completion of the exercise intervention), mainly due to missing second diaries, percentage missings for costs and EQ-5D scores were 16.5% and 18% at 32 weeks and 26% and 27% at 36 weeks, respectively.

We performed a sensitivity analysis excluding all patients with missing second diaries, which yielded similar results and did not change our conclusion:

For colon cancer, the incremental costs savings were €-5,450 (imputed dataset (IDS): €-4,305) and QALY improvements were 0.03 (IDS: 0.04) as compared to control. For breast cancer, incremental costs were €2,666 (IDS: €2,890) and the incremental effect was 0.01 QALY (IDS: 0.01)).

2) The manuscript says that "a bootstrap analysis" was used. However, the gold standard is BCa (bias-corrected and accelerated) bootstrap. Was this done here, and if so, please state that, and if not, why not?

Authors' answer:

We used the standard (percentile-t) bootstrap method. Within a CEA context, the proposed BCa (bias-corrected and accelerated) bootstrap method is not always needed and will most likely not change results in situations with a large standard error of the ICER (Briggs, 1997; Campbell 1999). However, we realized that a 1000 sample bootstrap is relatively small and therefore increased the bootstrap to 5000 samples (as is now also indicated in the statistical analyses section). This led to small changes in the numeric results, but not in our conclusions. We changed all numbers in the manuscript accordingly.

References:

Briggs AH1, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. *Health Econ.* 1997 Jul-Aug;6(4):327-40.
Campbell MK, Torgerson DJ. Bootstrapping: estimating confidence intervals for cost-effectiveness ratios. *QJM.* 1999 Mar;92(3):177-82.

3) On page 10 in the section labeled QALYs, all results are reported as a change in the number of QALYs. However, Table 3 shows the first two of these values for each type of cancer as "Average EQ-5D" scores, and from the results shown there it looks like that is what they are. The QALY values that are labeled as such in Table 3 are reported correctly in the text.

Authors' answer:

Indeed, in Table 3 the first two columns show the "Average EQ-5D" scores per group. In the third column the mean between-group differences between the "Average EQ-5D" scores are displayed. These between-group differences are reported on page 10 (and not changes over time).

4) On Table 3 both the EQ-5D score and the EQ-5D VAS are reported. I would recommend that the EQ-5D VAS values be removed from the table because they are inappropriate for use in a societal perspective analysis. The EQ-5D VAS measures HRQoL from the patient's perspective, whereas, the

EQ-5D has (Danish) societal preference weights applied to give HRQoL for the reported health states from the societal perspective. I would also recommend removal of any mention of the EQ-5D VAS from the text unless the authors want to explain its use.

Authors' answer:

We agree with the reviewers and removed any reference to the EQ-5D VAS from the table and the text.

5) Also, taking the EQ-5D VAS numbers out of Table 3 would give more room to report the averages of the full set of 9 every-4-week EQ-5D scores. Given the dramatic differences seen between the BCA and CCA cohort, providing more transparency here would help. I also wonder if the authors applied the full AUC approach as recommended by Manca (Health Econ. 14: 487–496 (2005)) including the adjustment for differences in baseline values. If not, that is recommended.

Authors' answer:

We added the average EQ-5D scores as assessed every 4 weeks in Table 3.

We thank the reviewer for recommending the full AUC approach. We now used the regression based method with correction for baseline values as proposed by Manca (and we added Manca et al to our references). The adjusted QALY totals over 36 weeks are added to Table 3. We used the adjusted QALYs as effect measure in all cost-effectiveness calculations.

Using this approach led to a small change in numeric results for colon cancer and a neglectable change for breast cancer. Our conclusions remain the same.

6) On the bottom of page 13 a statement is made that "about 50% of control group patients also reported high levels of physical activity." Given the dramatic differences seen between the QALY results for the BCA and CCA cohorts, it is important to indicate whether both these groups experienced this problem equally.

Authors' answer:

Contamination was not different between breast and colon cancer patients. In order to clarify this, we changed the sentence at the bottom of page 13 as follows:

"First, due to the nature of the intervention the patients were not blinded and about 50% of control group breast and colon cancer patients also reported high levels of physical activity,..."

VERSION 2 – REVIEW

REVIEWER	Patricia Herman RAND Corporation Santa Monica, CA, USA
REVIEW RETURNED	19-Oct-2016

GENERAL COMMENTS	<p>Thank you for your responses to my previous comments. Minor revisions remain.</p> <p>Please move the two sentences regarding the baseline adjustment of QALYs from the end of the Statistical Analysis section to the Utilities section after the sentence on the AUC method. It is confusing to have these highly related statements so separated.</p>
-------------------------	--

	<p>The statement in the discussion section about the control group patients reporting high levels of physical activity to could be made even clearer by adding the word "both" before the word "control" - "about 50% of both control group..."</p> <p>Table 3. Thank you for adding the every 4 week EQ-5D scores. Very enlightening. It looks like commas were accidently substituted for decimal points in both the adjusted QALY totals over 36 weeks. Also, the reporting of these results in the text remains problematic. The text in the section titled QALYs starting on the bottom of page 9 reports the mean difference in the average EQ5D scores during each 18 week period from Table 3 (i.e., the 0.001 and the 0.02 for breast cancer at the bottom of page 9 and the 0.02 and 0.11 for colon cancer at the top of page 10) as QALYs when they are not. These numbers would have to be adjusted for the time period they represent to be QALYs--i.e., each multiplied by 18/52. Best would be to call them what they are, differences in average EQ-5D scores during each of those 18 week periods as shown on Table 3.</p> <p>Minor typos: Page 5, last sentence of Exercise Intervention paragraph - "has" should be "have" Page 10, second line, the word "adjustment" is mis-spelled</p>
--	--

VERSION 2 – AUTHOR RESPONSE

Comments reviewer 2

1) Please move the two sentences regarding the baseline adjustment of QALYs from the end of the Statistical Analysis section to the Utilities section after the sentence on the AUC method. It is confusing to have these highly related statements so separated.

Authors' answer:

We moved the two sentences to the Utilities section.

2) The statement in the discussion section about the control group patients reporting high levels of physical activity to could be made even clearer by adding the word "both" before the word "control" - "about 50% of both control group..."

Authors' answer:

We added 'both' to the statement in the discussion.

3) Table 3. Thank you for adding the every 4 week EQ-5D scores. Very enlightening. It looks like commas were accidently substituted for decimal points in both the adjusted QALY totals over 36 weeks. Also, the reporting of these results in the text remains problematic. The text in the section titled QALYs starting on the bottom of page 9 reports the mean difference in the average EQ5D scores during each 18 week period from Table 3 (i.e., the 0.001 and the 0.02 for breast cancer at the bottom of page 9 and the 0.02 and 0.11 for colon cancer at the top of page 10) as QALYs when they are not. These numbers would have to be adjusted for the time period they represent to be QALYs--i.e., each multiplied by 18/52. Best would be to call them what they are, differences in average EQ-5D scores during each of those 18 week periods as shown on Table 3.

Authors' answer:

We changed the commas back to decimal points.

As suggested by the reviewer, we now use "differences in average EQ-5D scores" instead of "QALYs"

when reporting the results in the text.

Minor typos:

Page 5, last sentence of Exercise Intervention paragraph - "has" should be "have"

Page 10, second line, the word "adjustment" is mis-spelled

Authors' answer:

Thank you. We corrected both typos.