Reviewer’s report

Title: Can routine data be used to support cancer clinical trials? A historical baseline on which to build: retrospective linkage of data from the TACT (CRUK 01/001) breast cancer trial and the National Cancer Data Repository

Version: 0 Date: 19 Jan 2017

Reviewer: Richard Jackson

Reviewer’s report:

Overall this is a well organised and conducted manuscript answering a topical question in the field of trial design.

The manuscript is well structured and the results in the discussion/conclusion are supported by the data presents.

There are only a small number of minor comments which I feel may add to the research:

With respect to the multivariable logistic model - more detail (i.e. candidate covariates, selection methods) would add to the clarity

With respect to Overall Survival - Whilst there is very little difference in the two distributions presented in Figure 3 - there is some small evidence of divergence after 8 years. There might be some merit in performing landmark analysis (using landmarks of 5,6,7 & 8 years perhaps) to see if this divergence is more pronounced.

It is mentioned in the discussion that there was considerable amount of cleaning required of the data that was obtained - this raises a couple of questions. Firstly - if this approach was adopted in a future clinical trial, to what extent would the ‘cleaning’ have to be verified. Secondly, presumably there is considerable resources that go into the collection, cleaning and organisation of the data - is it possible to make any comment on the extent to which the resources incurred using this approach compare to standard data collection approaches?

Level of interest
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable
Quality of figures
All images and figures within the manuscript should be genuine i.e. without evidence of manipulation. No specific feature within an image may be enhanced, obscured, moved, removed, or introduced. If you have concerns about the veracity of the figures you should choose the first option below.

Statistical review
Is it essential that this manuscript is seen by an expert statistician? If so, please give your reasons in your report.

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

None

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.
I agree to the open peer review policy of the journal