Author's response to reviews

Title: Vocational rehabilitation services for patients with cancer: design of a feasibility study incorporating a pilot randomised controlled trial among women with breast cancer following surgery

Authors:

Richard G Kyle (richard.kyle@stir.ac.uk)
Bill Culbard (w.r.culbard1@stir.ac.uk)
Josie Evans (josie.evans@stir.ac.uk)
Nicola M Gray (n.gray@abdn.ac.uk)
Dolapo Ayansina (d.ayansina@abdn.ac.uk)
Gill Hubbard (gill.hubbard@stir.ac.uk)

Version: 3 Date: 17 March 2011

Author's response to reviews: see over
Dear Editor

Re: MS: 9261582944669231

Vocational rehabilitation services for patients with cancer: design of a feasibility study incorporating a pilot randomised controlled trial among women with breast cancer following surgery

Thank you for providing the opportunity to respond to reviewer comments. We have addressed their comments accordingly:

<table>
<thead>
<tr>
<th>Comment</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. P(age) 1, p(aragraph) 2, l(ine) 3. Suggest dropping [old].</td>
<td>[Old] has been deleted.</td>
</tr>
<tr>
<td>2. P 1, p 2, l 7. Suggest including the name of the QoL measure.</td>
<td>The names of the QoL measure [Fact-B] and fatigue measure [FACIT-Fatigue] have been inserted.</td>
</tr>
<tr>
<td>3. P 1, p 4, l 1. Include the date of registration and the date the first patient was randomised, if that has happened.</td>
<td>The registration date and the date the first patient was randomised has been added.</td>
</tr>
<tr>
<td>4. P 2, p 1, l 3. Insert [years] after [65]. Also, P 2, p 1, l 8.</td>
<td>[Years] has been inserted at both occurrences.</td>
</tr>
</tbody>
</table>
| 5. P 2, p 1, l 5. Since incidence is a time related rate, what is the time, is it [annual]; if so, insert it. | The sentence has been revised to clarify the time period: [Breast cancer is the most commonly diagnosed cancer in Scottish women and incidence has increased by 8% over the past decade attributed, in part, to earlier detection due to mammography screening programmes.]
| 6. P 2, p 1, l 8-9. Delete [old]. Also P 4, p 1, l 1. | [Old] has been deleted at both occurrences. |
| 7. P 2, p 3, l 2. Since [or] logically includes [and], drop [and]. | [and/] has been deleted. |
| 8. P 4, p 3, l 7. Rewrite as […] with 5% level of significance and …[]. It is not proper to use its compliment. | Rewritten as [with 5% level of significance]. |
| 9. P 5, p 1. There is no need to repeat the wording for those that are ineligible or refuse. Why not state it once? | The sentence has been rewritten to avoid repetition:
[In each hospital, patients who do not want to participate are asked if they consent to the retention of a limited set of anonymised clinical, sociodemographic and employment data (e.g.,] |
Re: MS: 9261582944669231

10. P 7, p 1, l 2. Replace [ranges] by [varies]. [ranges] has been replaced with [varies].

11. P 7, p 1, l 4. Replace [range] by [vary]. Also, P 7, p 2, l 5. [range] has been replaced with [vary] at both occurrences.

12. P 7, p 1. Did you consider adding something like the EQ-5D to get utilities as well as a disease specific measure? This would make the results easier to compare to other diseases other than Breast CA, and allow for a cost utility analysis after the main trial is completed. We thank the reviewer for this suggestion. Unfortunately, as the pilot RCT is currently in progress there is not an opportunity to incorporate this measure. However, if the study progresses to a larger more definitive trial we will consider inclusion of this measure to facilitate cost utility analysis. It should also be noted that the VR services to which women in the intervention arm are referred use this measure at baseline and 3- and 6-months follow up and these data will be obtained from the service.

13. P 7, p 5, l 1. Provide a citation for why you think 5 patients per group for a total of 10 will be adequate for the qualitative analysis. Similarly P 8, p 2. This has been addressed as part of Referee 4’s fourth point (please see below).

14. P 8, p 3, l 2,3. Provide references to the two software packages. PASW and NVivo have now been referenced.

15. P 9, p 1, l 1. There is no need to test for comparison of the two groups after they are randomized. It is much more important to measure the randomization integrity. Comparison should be done clinically as all the differences are expected to be due to chance. See Altman DG & Dore CJ: Randomisation and baseline comparisons in clinical trials. Lancet 1990;335(8682):149-53 and a letter in Lancet also 1990;335(8703):1476 and another paper by them: Baseline comparisons in randomized clinical trials. Stat Med 1991;10(5):797-9. One needs to demonstrate integrity of the randomizations, not test the baseline variables. It is not our intention to test for comparison of the two groups after they are randomised, and we apologise for the lack of clarity in the text which may have led to this interpretation. Rather, in accordance with current convention, we will report baseline characteristics as mean and standard deviation for continuous data and n (%) for categorical data. We have revised the ‘Statistics’ section to clarify this point.

A random sample of 10 references was checked for accuracy of citation. 16. References 1,8,9,10,13,19 seem accurate. No action required.

17. P 10, R 4 should have location of an organization to help find it. This reviewer could not locate it. Is this part of the Scottish government? The organisation [NHS National Services Scotland] has been added.

18. P 10, R 7, l 2. Insert [(6)] after [14]. [(6)] has been inserted after [14].

19. P 10, R 1 1. Delete [.] after the initial of the first author. [.] has been deleted.

20. P 11, R 12, l 2. Insert [(7)] after [173]. [(7)] has been inserted after [173].


and l 2 insert [(2)] after [51]. [(2)] has been inserted after [51].

Referee 2
First, I have grave doubts about the sensitivity of the FACT questionnaire to effects resulting from the intervention. It seems likely that patients who return to work earlier due to vocational rehabilitation will age, sex, postcode, diagnosis, treatment, employment status) so that a comparison can be made between patients who do and do not agree to be approached (Figure 1).] This potential limitation of the FACT measure has been acknowledged by revising the following paragraph:

[As this is a feasibility study, hypothesis testing...]

Page 2 of 6
have a better quality of life than women who don’t, but I doubt whether the FACT will pick it up. The authors are advised to read each and every question on the FACT instrument and ask themselves whether they believe that the intervention will lead to a large difference in responses (small differences are very hard to pick up). For example, is it plausible that vocational rehabilitation will lead to detectable differences in responses to items such as “I feel close to my friends” or “I have difficulty in eating heavy foods”. At the very least, the authors might consider adding evaluation of the outcome questionnaires as a study aim. The authors do list as an objective determining whether the outcome measures are “appropriate”, but do not state how this will be done.

Second, the statistical analysis section is inadequate. Yes, this is only a feasibility study, but that doesn’t mean “anything goes”. In particular, even if just a feasibility study, the authors will undoubtedly report an estimate of treatment effect, and the statistical methods used need to be specified in the protocol so that they can be compared with those reported in the trial report. For example, the authors state “The differences from baseline values to 6 month and 12 month follow-up for primary and secondary outcome measures will be compared using appropriate paired and independent sample tests”. This is extremely vague! The protocol for comparison of outcomes between groups should be precise enough so that two competent statisticians reading the protocol would conduct identical analyses and report identical statistical results. As a general word of advice: use of paired tests is irrational for a randomized trial; outcomes for which baseline and follow-up measures are taken should be analyzed by ANCOVA (see Vickers Altman BMJ);

the main endpoint is a time-to-event variable and so should be analyzed using Cox regression.

On this point: there doesn’t appear to be a statistician listed as an author; might it be wise for a statistician to be included on the study team?

must proceed cautiously. However, it is anticipated that participants referred to VR services will experience fewer days off work due to sickness, less fatigue and increased QoL. QoL measures will be analysed at global (i.e., summed FACT-G and FACT-B scores), sub-scale (i.e., physical, social/family, emotional, BCS), and individual question level. This will also enable evaluation of the outcome questionnaire’s ability to detect differences in QoL between the intervention and control group associated with VR intervention. The greatest difference may be detectable at the level of individual domains of well-being (particularly, emotional, social/family and functional) and individual questions included in these domains (e.g., “I am satisfied with how I am coping with my illness” [GE2], “I feel close to my friends” [GS1], “I get support from my friends” [GS3], “I am able to work (include work at home)” [GF1], “My work (include work at home) is fulfilling” [GF2]) as support to remain in or return to employment may increase emotional well-being, enable access to a wider network of social support through colleagues, and increase physical functioning.

We apologise for the lack of detail in the statistical analysis section. We have revised this to specify the use of an independent samples t-test and ANCOVA to statistically test differences in the primary and secondary outcome measures respectively between the control and intervention groups.

Since we are not measuring a time-to-event (such as days to return to work) but, rather, comparing the number of days of sick leave between the intervention and control group we believe the use of Cox regression is inappropriate in this case.

A statistician (DA) is part of the study team and is responsible for providing guidance on statistical analysis of trial data. He is now listed as a co-author on this paper.
Third, the investigators state as an aim that “response rates” will be evaluated. It is unclear what is meant by this term.

Clarification of the meaning of this term within the pilot RCT has been added: [Estimate trial recruitment rates (i.e., percentage of eligible women who consent to the trial)...].

Referee 3

Too many aims
Nine is too many aims for most studies, especially a pilot study. Some of the aims, such as aims 1 and 3, both of which address determining the number of potentially eligible patients, could be combined into a single aim. Others belong in a different study, such as aim 8, which addresses the suitability of patients with other tumor types for a study of vocational rehab and return to work.

The number of aims has been reduced to 6 through both combination (of aims 1 and 3, and 4 and 5) and deletion (of aim 9, which refers to 12-month follow-up which is being conducted as part of the pilot RCT). The revised aims are now the following:

1. Determine the numbers of patients in Lothian and Tayside with breast cancer referred for surgery who are in paid or self-employment and potentially eligible for the pilot RCT.
2. Assess whether it is feasible and acceptable to recruit patients into the trial post-operatively. If not, identify the most appropriate timing for recruitment and the employment intervention that follows.
3. Estimate trial recruitment rates (i.e., percentage of eligible women who consent to the trial) and attrition over the period of the trial.
4. Determine whether the outcomes are measurable and appropriate, and evaluate the instruments used to measure secondary outcomes (i.e., FACT-B, FACIT-Fatigue).
5. Estimate the likely effect size to inform the power calculation for a larger, more definitive trial.
6. Assess the feasibility of also including patients with lung and prostate cancer in a larger future RCT.

We believe that aim 6 (originally aim 8) is an integral part of this study since assessing the feasibility of the trial and acceptability of the intervention among people with prostate and lung cancer will be key to the successful development of both a larger more definitive future trial and the vocational rehabilitation interventions this pilot RCT seeks to inform.

Unnecessarily complex design
It is unclear to me why a feasibility study needs to be designed as a randomized trial. The randomized design does nothing to advance any of the stated aims. If the overall goal is really to assess the feasibility of undertaking a larger study designed to assess the impact of a vocational rehabilitation intervention on return to work after cancer, then the pilot could be done as a single arm, with all participants receiving the intervention. The only aim that could benefit from a randomized design is aim 7, which addresses estimating an effect size for the purposes of power calculations for the future larger study. However, the investigators could estimate the effect size to inform sample size calculations by looking at some of the published studies

We have inserted the following sentence to justify the selection of this study design in the ‘Background’:

[A study design incorporating a pilot RCT was selected to refine and assess the feasibility of trial processes (including recruitment, randomisation and follow-up) and estimate the likely effect size in advance of a larger, more definitive, future trial.]

Although complex, we believe that the best way to assess the feasibility of a larger trial is to run this pilot. For example, we currently do not have data on eligibility for VR interventions among this (or indeed other cancer) population(s), principally because employment data are not routinely collected. Furthermore, running a pilot RCT enables recruitment processes to be ‘road-tested’ in each site to ensure seamless integration with...
addressing vocational rehabilitation and return to work.

With respect to the sample size calculation for the pilot study, I do not understand how the authors arrived at a standard deviation of 100 days. Is this taken from the study by Drolet et al. that is cited in the preceding sentence? This should be stated explicitly.

<table>
<thead>
<tr>
<th>Lack of standardization of the intervention</th>
<th>This paragraph has been revised to provide readers with further information about this complex intervention to address the reviewer’s concerns as follows:</th>
</tr>
</thead>
<tbody>
<tr>
<td>The authors state that “the intervention is referral to a VR service.” Will all participants be referred to the same vocational rehabilitation provider? Will there be any standardization of the intervention? It seems that someone not involved in the study (the “case manager”) will determine the kind of services each participant will receive, but it is not at all clear that this person will follow a predetermined standardized system to assign services. On what will they base these important decisions? How will the investigators keep track of the specific intervention that is received by each participant? All of this needs to be explained clearly. If the authors plan to study an intervention that is not standardized, then they should explain their rationale for doing so.</td>
<td></td>
</tr>
<tr>
<td><strong>Lack of standardization of the intervention</strong> The authors state that “the intervention is referral to a VR service.” Will all participants be referred to the same vocational rehabilitation provider? Will there be any standardization of the intervention? It seems that someone not involved in the study (the “case manager”) will determine the kind of services each participant will receive, but it is not at all clear that this person will follow a predetermined standardized system to assign services. On what will they base these important decisions? How will the investigators keep track of the specific intervention that is received by each participant? All of this needs to be explained clearly. If the authors plan to study an intervention that is not standardized, then they should explain their rationale for doing so.</td>
<td>This paragraph has been revised to provide readers with further information about this complex intervention to address the reviewer’s concerns as follows:</td>
</tr>
<tr>
<td><strong>The complex intervention is referral to a VR service in either Tayside or Lothian.</strong> Patients recruited from PRI and Ninewells, Dundee are randomised to receive referral to a VR service in Tayside; individuals enrolled from WGH are referred to a VR service in Lothian. All participants allocated to the intervention arm of the trial are contacted by a VR service by telephone within 10 days following return of the baseline questionnaire to the researcher (RGK). Participants are allocated a ‘case manager’ who conducts a telephone assessment of supportive care needs to facilitate remaining in or returning to work during which baseline measures are also recorded using the Canadian Occupational Performance Measure (COPM)[13], 12 item General Health Questionnaire (GHQ-12)[14] and European Quality of Life – 5 Dimensions (EQ-5D)[15]. Outcomes are re-measured at 3- and 6-months from referral. Based on this initial assessment of each individual’s personal goals and health status the case manager signposts participants to appropriate support services including physiotherapy, occupational therapy, occupational health nurse, occupational health doctor, counsellor/psychological therapy, complementary therapy. Each individual may therefore receive a different (combination of) intervention(s). Usual care following surgery involves no formal employment support. Participants in both arms of the trial receive a copy of the booklet Work and Cancer published by Macmillan Cancer Support.[16]</td>
<td></td>
</tr>
<tr>
<td>As the personalised and complex nature of the intervention precludes exogenous standardisation, data will be obtained from each service on the specific interventions received by individuals in the intervention arm of the pilot RCT. Secondary analysis of outcome measures recorded by each service at initial assessment and 3- and 6-month follow up will also be undertaken for these women. In addition, information will be gathered about each service...</td>
<td>This figure is an estimate derived following consultation with a statistician. We have revised the text accordingly to explicitly state this figure is an estimate.</td>
</tr>
</tbody>
</table>

---

Re: MS: 9261582944669231

Clinical practice. This will ensure that a larger, more definitive, future trial runs as smoothly as possible from the outset. Even in the short time since recruitment began the rigour demanded by an RCT in terms data collection and recruitment procedures has laid a solid research foundation in each clinical team upon which a future study can build.
Qualitative methodology
When undertaken correctly, qualitative research can be extremely valuable and informative, particularly when undertaking a program of research in which there is scant existing data. The semi-structured interview format is appropriate for this study, but these interviews should be done until a point of saturation is reached. The decision to interview 5 participants in the treatment arm and 5 in the control arm seems arbitrary. Instead, the researchers should interview 3-5 in each arm and then analyze the data, with a plan to do additional rounds of interviews (of new participants) until no significant additional information is learned from the interviews.

We welcome this reviewer’s comment on the qualitative methodology. It is our intention to adopt this approach and increase the number of interviewees until no new themes emerge. However, we neglected to make this clear, for which we apologise. In order to provide clarification we have inserted the following sentence in the paragraph describing the interviews with trial participants:

Maximum variation sampling will be used to initially identify 5 patients in each arm of the trial. Key variables in the sampling frame will include: current employment status (i.e., at/off work), job, role, tenure (i.e., part-/full-time), age, diagnosis, treatment pathway. Interviewing and analysis will proceed concurrently to pinpoint data saturation, and further rounds of interviews with new participants will be conducted, if required.

This is also the case for the interviews of lung and prostate cancer patients,

although, as I explained above, I believe this part of the study should be undertaken separately (see “too many aims” above).

We have also added the following to the paragraph discussing the interviews with people with lung and prostate cancer:

Maximum variation sampling will be used to initially identify 5 patients with lung cancer and 5 people with prostate cancer. Key variables in the sampling frame will include: current employment status (i.e., at/off work), job, role, tenure (i.e., part-/full-time), age, diagnosis, treatment pathway. Individuals who were first treated with surgery will be approached to enable comparison with the group of women in the pilot RCT. Interviewing and analysis will proceed concurrently to pinpoint data saturation, and further rounds of interviews will be conducted with new participants, if required.

The names of public figures should be stated completely (see the second paragraph of page 2, “Dame Black’s recent review of the health . . .”)

Dame Black’s forename has been added: [Dame Carol Black’s recent review of the health of Britain’s working age population...].

In the last paragraph of page 4 the researchers alternate between the future and present tenses. This needs to be revised to be consistent throughout the paragraph (and the entire section describing the study methods).

The tense of this paragraph, in particular, and the ‘Methods’ section, in general, has been revised to reflect the fact that the pilot RCT is currently recruiting participants.

I have today uploaded a revised version of the manuscript indicating these tracked changes.

I look forward to hearing the outcome of your deliberations.

Yours sincerely,

Dr Richard G Kyle