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)	How different terminology for ductal carcinoma in situ impacts women's concern and treatment preferences: a randomised comparison within a national community survey
AUTHORS	McCaffery, Kirsten; Nickel, Brooke; Moynihan, Ray; Hersch, Jolyn; Teixeira-Pinto, Armando; Irwig, Les; Barratt, Alexandra

VERSION 1 - REVIEW

REVIEWER	Kim Sprunck Dana-Farber Cancer Institute USA
REVIEW RETURNED	06-May-2015

GENERAL COMMENTS	Overall, I think that the intent of the content of this article is interesting. However, I did not find the conclusions to be in sync with what the authors propose. Below are some specific pointers about the paper. Abstract: 1) it would be helpful to note in the abstract settings/participants that some of the women had been breast cancer affected. Moreover, actually, I think best to take affected women out of the sample as they are too exposed to the topic at hand. 2) In the results section: a. overall I thought the results section could be written more clearly. While was able to follow it, it was a bit confusing overall. b. please clarify from what did arm A changed their preference from treatment (from watchful waiting to treatment?). c. the conclusion says 'carcinoma' should not be used...though I did not see that this word was actually ever used in your study? Please clarify. There were no limitations stated in the strengths and limitations section. Every study has limitations. Main Paper Introduction Since this was an Australian based study, why did you use US stats of DCIS? Please clarify and/or at least add some Australian stats. Methods: Please explain why men were asked these questions, but then later excluded. If you did not need the data, why did you include this group in the questions asked? Results: The most overarching concern that I have with this paper is that I do

	not think the conclusions are correct. The authors state that the study was set to "examine the effect of different terminologies used to describe DCIS, with and without the word cancer, on hypothetical concern...". The authors did do this, with randomizing the groups at the beginning and asking these 'clean' groups their thoughts with the differing language between arm A and arm B. The authors found that "initial concern was high in both arms..."...I believe that is the conclusion. There was no difference, no matter the wording used. When looking at time two, because groups had already been exposed to previous language, I am not sure you can come to the conclusion that the authors did in their paper that "our research suggest that changing terminology impacts women's concern..." Again, because of the randomization a the beginning...those would be your comparative groups, and there was no difference. This holds true for treatment preferences. The authors state that "preferences for watchful waiting were high." So, again, these groups show that no matter the language used, watchful waiting was desired. In order for this paper to move forward, I think that the paper would need to be restructured and written with these themes central.

REVIEWER	Jo Waller University College London, UK
REVIEW RETURNED	17-Jun-2015

GENERAL COMMENTS	This is a well-written and interesting paper seeking to explore the impact of different terminology on women's responses to a diagnosis of DCIS. The study is clearly set in context, the methods and results are well described and the discussion puts the findings in the context of the (scant) existing literature. I'm not sure I completely agree with the conclusions that the authors draw, though. There main finding was that there was no initial difference between the two groups of women given different terminology. I think this is a clear finding and one that is rather played down in the discussion and conclusions. The finding that women change their responses when presented with the alternative terminology is interesting, but I think it's harder to interpret. The issue of demand characteristics is not discussed - it seems possible that if women are given slightly different information and then asked 'And now what do you think?', some may feel that they ought to change their response. The differential change between the two groups provides some reassurance that this is not the case, but I think it should be discussed as a possibility. I believe this is an important paper, and opens up a new area for future research, but I would suggest that the conclusions should be a little more tentative, given that if the two terms are really interpreted differently, there should have been an initial between-group difference.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1.

1. Abstract:
I think best to take affected women out of the sample as they are too exposed to the topic at hand.

RESPONSE: We have removed affected women out of the analyses.

2. In the results section: a) overall I thought the results section could be written more clearly. While was able to follow it, it was a bit confusing overall. b) please clarify from what did arm A changed their preference from treatment (from watchful waiting to treatment?).

RESPONSE: We have redrafted the results section of the abstract as follows to make it clearer:

"Women in both arms indicated high concern, but still indicated strong initial preferences for watchful waiting (64%). There were no differences in initial concern or preferences by trial arm. However, more women in arm A ("abnormal cells" first term) indicated they would feel more concerned if given the alternative term ("pre-invasive breast cancer cells") compared to women in arm B who received the terms in the opposite order (67% arm A vs 52% arm B would feel more concerned, p=0.001). More women in arm A also changed their preference towards treatment when the terminology was switched from "abnormal cells" to "pre-invasive breast cancer cells" compared to arm B. In arm A 18% of women changed their preference to treatment while only 6% changed to watchful waiting (p=0.008). In contrast, there were no significant changes in treatment preference in arm B when the terminology was switched (9% vs 8% changed their stated preference)."

We have also added an additional table in Appendix 2 to further clarify our results. This is referenced on page 9 of the manuscript in the results.

c) the conclusion says 'carcinoma' should not be used...though I did not see that this word was actually ever used in your study? Please clarify.

RESPONSE: We have clarified this section as follows: "Removal of the cancer term from DCIS may assist in efforts towards reducing overtreatment".

d) There were no limitations stated in the strengths and limitations section. Every study has limitations.

RESPONSE: We have added the following two limitations to this section on page 3
• Limitations of the study include its hypothetical design, therefore women facing a real diagnosis of DCIS may respond differently.
• Preferences for watchful waiting were predicated on the statement, "if research shows watchful waiting is a safe and effective option"; the subject of two current randomised trials of women with DCIS.

3. Introduction
Since this was an Australian based study, why did you use US stats of DCIS? Please clarify and/or at least add some Australian stats.

RESPONSE: These are now added on page 4.

4. Methods: Please explain why men were asked these questions, but then later excluded. If you did not need the data, why did you include this group in the questions asked?

RESPONSE: This paper concerns women's responses to breast cancer terminology therefore only women are included.

5. Results: a) The most overarching concern that I have with this paper is that I do not think the conclusions are correct. The authors state that the study was set to "examine the effect of different terminologies used to describe DCIS, with and without the word cancer, on hypothetical concern...". The authors did do this, with randomizing the groups at the beginning and asking these 'clean' groups their thoughts with the differing language between arm A and arm B. The authors found that "initial concern was high in both arms..."...I believe that is the conclusion. There was no difference, no matter the wording used. When looking at time two, because groups had already been exposed to previous language, I am not sure you can come to the conclusion that the authors did in their paper that "our research suggest that changing terminology impacts women's concern..."

RESPONSE: We have modified the conclusions of our study to address the reviewers concerns to further highlight there was no significant difference in the initial responses given by arms A and B. However we believe we can compare the impact of switching terminology in our study which randomised women in arms A and B to receive terminology in the opposite order and assessed change in concern and treatment preferences. We are now more precise in our aims and our interpretation of our findings in several places in the manuscript:

Abstract conclusion page 2
"In a hypothetical scenario interest in watchful waiting for DCIS was high and changing terminology impacted women's concern and treatment preferences. Removal of the cancer term from DCIS may assist in efforts toward reducing overtreatment".

Introduction page 5
"This study set out to examine whether the use of terminology including the term cancer to describe DCIS, increased hypothetical concern and treatment preferences among a community sample of Australian women.

Discussion page 10
"In a randomised comparison of terms for DCIS among a national community sample of women, interest in watchful waiting was high irrespective of the terminology used".

Discussion page 12
"Our research suggests that switching terminology from "abnormal cells" to "pre-invasive breast cancer" impacts women's concern and treatment preferences at least in a hypothetical setting",

b) Again, because of the randomization at the beginning...those would be your comparative groups, and there was no difference. This holds true for treatment preferences. The authors state that "preferences for watchful waiting were high." So, again, these groups show that no matter the language used, watchful waiting was desired.

RESPONSE: Addressed above

Reviewer 2
The study is clearly set in context, the methods and results are well described and the discussion puts the findings in the context of the (scant) existing literature.

RESPONSE: Thank you

I'm not sure I completely agree with the conclusions that the authors draw, though. There main finding was that there was no initial difference between the two groups of women given different terminology. I think this is a clear finding and one that is rather played down in the discussion and conclusions.

RESPONSE: Addressed in several places in the manuscript as described above.

The finding that women change their responses when presented with the alternative terminology is interesting, but I think it's harder to interpret. The issue of demand characteristics is not discussed - it seems possible that if women are given slightly different information and then asked 'And now what do you think?', some may feel that they ought to change their response. The differential change between the two groups provides some reassurance that this is not the case, but I think it should be discussed as a possibility.

RESPONSE: We understand the reviewers concerns however, since women in the randomised arms show a statistically significant difference in the change in each arm which is consistent across both outcomes (concern and treatment preference with both increased concern and increased treatment preference shown in arm A), we feel that the results cannot simply be a result of demand characteristics. We have added this point to the discussion page 12:

“We also note that a significant difference was observed in arm A when terms were switched from non-cancer to cancer. Although it is possible, we do not believe this difference was a result of demand characteristics from the study design (with women feeling obliged to change their response). If this was the case the same pattern would be expected in both randomised arms but was not observed in the arm B. In addition, the changes we observed in arm A were in the same direction and consistent for both outcomes, concern and treatment preference”.

I believe this is an important paper, and opens up a new area for future research, but I would suggest that the conclusions should be a little more tentative, given that if the two terms are really interpreted differently, there should have been an initial between-group difference.

RESPONSE: We have changed our discussion and conclusions to reflect a more tentative approach to our findings (See abstract, Discussion page 10 and 12).

VERSION 2 – REVIEW

REVIEWER	Kim Sprunck Dana-Farber Cancer Institute, USA
REVIEW RETURNED	11-Sep-2015

GENERAL COMMENTS	The authors have done a nice job of addressing the major concerns previously documented. There are just a couple of minor issues and one overarching one. Minor issues: 1) education level statement confusion-- in the results section it is stated that that sample had slightly higher levels of education than that of the general Australian population. on pg 10, line 48 and pg 11, line 38 it says that your population had a high proportion from lower educational backgrounds. Both of these statements could be true, but seem contradictory.

	2) there is a type o, page 10, line 40, 'an a difference' 3) While not the full intent of this paper, the authors point out that part of the desire to change terminology is to potential reduce over diagnosis and over treatment of DCIS. they do include one sentence at the top of pate 12 that doctors need to support watchful waiting. Clearly changing terminology said to patients may help with the reduction of over treatment, but it really is in the hands of the medical field to agree to watchful waiting. In the discussion section, it would be helpful to add a couple sentences about any work that has been done with doctors about why treatment is so often pursued vs. watchful waiting. Overarching comment: As noted, the authors definitely addressed the previous concerns. So, on that front, they have fulfilled the requests for changes. However, I think I'm now left with a question of what does this paper add to the literature? Since women in both groups had high initial concern and strong initial preferences for watchful waiting, understanding that there was a statistically significant difference in strength of concern for Arm A to Arm B women, I'm not sure about the usefulness about knowing concern changes with terminology differences with each individual, as this is not how it (changing the explanation/description of a DCIS dx) would happen in the real world.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
The authors have done a nice job of addressing the major concerns previously documented. There are just a couple of minor issues and one overarching one.

Minor issues:
1. Education level statement confusion-- in the results section it is stated that that sample had slightly higher levels of education than that of the general Australian population. on pg 10, line 48 and pg 11, line 38 it says that your population had a high proportion from lower educational backgrounds. Both of these statements could be true, but seem contradictory.

We have added in a clarifying sentence in the results and discussion to remove any confusion about the sample's education level. Please see page 8, 11 and table 1 and below.

Page 8:
'Although the sample included a large proportion of women with low levels of education (48%), overall the women in our study had slightly higher levels of education than that of the general Australian population.'

Page 11:
'First, our sample was a national community sample of women which included a large proportion of women with lower levels of education (48%), i.e. either did not complete high school or had no post-high school qualifications (age 18 years).'

2. There is a type o, page 10, line 40, 'an a difference'

The correction has now been made. Please see page 10.

3. While not the full intent of this paper, the authors point out that part of the desire to change terminology is to potential reduce over diagnosis and over treatment of DCIS. they do include one sentence at the top of page 12 that doctors need to support watchful waiting. Clearly changing terminology said to patients may help with the reduction of over treatment, but it really is in the hands of the medical field to agree to watchful waiting. In the discussion section, it would be helpful to add a couple sentences about any work that has been done with doctors about why treatment is so often pursued vs. watchful waiting.

We have now added in an additional sentence to explain why clinicians may not currently be willing to accept watchful waiting as a management option for DCIS. Please see page 12 and below.

'It is currently extremely difficult to test different terminologies among women diagnosed with DCIS (outside a clinical trial of watchful waiting for DCIS) since this needs to be a treatment option supported by clinicians. This may not be the case until rates of progression of DCIS to invasive cancer and the impact of watchful waiting are better established and understood through randomised controlled trials of this management strategy.23 24'

Overarching comment:
As noted, the authors definitely addressed the previous concerns. So, on that front, they have fulfilled the requests for changes. However, I think I'm now left with a question of what does this paper add to the literature? Since women in both groups had high initial concern and strong initial preferences for watchful waiting, understanding that there was a statistically significant difference in strength of concern for Arm A to Arm B women, I'm not sure about the usefulness about knowing concern changes with terminology differences with each individual, as this is not how it (changing the explanation/description of a DCIS dx) would happen in the real world.

Although there was not a difference in level of concern initially (as both groups indicated they had high concern), when the terminology was switched within the same individual from a non-cancer to a cancer term participants indicated a higher level of concern and a higher preference for immediate treatments. Although, as mentioned the terminology would not be switched in the real world, this indicates that among a national community sample of women hearing the term cancer elicited more concern and a higher preference for treatment which demonstrates that the terminology used during a diagnosis may change women's psychological responses and behaviour. We have now adjusted our final paragraph on pages 12 & 13 to make our conclusions about what our paper adds to the literature clearer.

Pages 12 & 13:
'There is growing concern about the problem of overdiagnosis and overtreatment of inconsequential disease. One strategy to mitigate this problem may be to change the terminology currently used to describe cancer related conditions which have low malignant potential,15 such as DCIS. This could potentially encourage both clinicians and patients to opt for more conservative treatment strategies such as watchful waiting although it would have to be part of a broader effort to support conservative treatment. Our research shows that a national sample of women demonstrated high levels of interest in watchful waiting, which has not previously been reported and supports the need for the two trials currently underway on this topic.23 24 Switching terminology from "abnormal cells" to "pre-invasive breast cancer" influenced women's concern and treatment preferences at least in a hypothetical setting. Together, the findings provide evidence that further investigation of the effects of changing DCIS terminology is needed in clinical populations as removing the cancer term may reduce concern and overtreatment, as proposed by Esserman et al.15 At minimum it shows that language is a powerful tool that has the potential to shape both understanding and actions.'