

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)	Should women aged 70-74 be invited to participate in screening mammography? A report on two Australian community juries
AUTHORS	Degeling, Chris; Barratt, Alexandra; Aranda, Sanchia; Bell, Robin; Doust, Jenny; Houssami, Nehmat; Hersch, Jolyn; Sakowsky, Ruben; Entwistle, Vikki; Carter, Stacy

VERSION 1 – REVIEW

REVIEWER	Nancy Schoenborn Johns Hopkins University School of Medicine, USA
REVIEW RETURNED	21-Dec-2017

GENERAL COMMENTS	<p>Overall comments: The authors do a wonderful job giving a clear account of their project which adds important knowledge to the literature. Using community juries, the authors explored the decision whether mammogram screening invitation to women ages 70-74 should continue. The findings are presented clearly followed by thoughtful discussions. I have only minor comments.</p> <ol style="list-style-type: none">1. Background: the rationale for conducting the study is well-laid out. Although I agree with the authors that the community juries method is a valuable approach to address the study question, it would strengthen the paper if there is available literature from survey studies of the broader Australian public on their opinions regarding invitation of older women to mammography screening. I would have liked to see a baseline survey of the jurors opinions before any evidence is presented, as this was not done in this study, having a baseline description (if available) of what the public thinks at baseline would present helpful comparison/contrast with the study results.2. Methods: I applaud the authors for presenting the method of community juries in a way that's quite clear to those who may not be familiar with the approach. I wonder if the authors can add a sentence on why two juries (as opposed to three for example) – is it the norm? is it partially because the same verdict was reached in both juries?3. Discussion: I believe that it's worthwhile to mention under limitations that the findings are also influenced by what specific information were presented in the juries, how they were presented, and by which specific experts. For example in retrospect, perhaps a clinician or even a patient instead of an epidemiologist who can comment on more personal aspects of the harms of overdiagnosis and overtreatment may have had more sway with the juries.4. Discussion: Related to the point above (#3), I wonder if the authors may be able to comment on how the information is presented to the juries may have impacted their opinions. It seems like the juries did not resonate with information that are population-
-------------------------	---

	based and value more personal and symbolic meanings around screening. I wonder if, in future studies, presenting for example the harms of screening with more personal stories and examples may lead to different responses.
--	--

REVIEWER	Lucie Rychetnik School of Medicine Sydney, University of Notre Dame, Australia
REVIEW RETURNED	15-Jan-2018

GENERAL COMMENTS	<p>This is an excellent paper that presents important findings that are highly relevant to current breast cancer screening policy and practice. I have provided below some comments / queries relating to specific sections of text for consideration by the authors prior to publication.</p> <p>P4, line 53 'they provide evidence' It would be useful to elaborate / distinguish between epidemiological evidence and the type offered by the findings of community juries – how does it assist policy making.</p> <p>P5, line 5 'all stake holders agreed' Why did all stakeholders all agree this was the most important question to address? Was a change of policy being considered? Was a change in current policy being advocated or did some of the groups want to advocate this? Also did the stakeholders and researchers discuss in advance how the findings of this study would be used? It would be useful to include here some more information about the policy context in which this research was being conducted.</p> <p>P5, line 51 'interrogating the evidence' Would be good to include some more detail of what type of evidence was interrogated (assume this means epidemiological or also other?). And the 'clinical and practical issues' ... were these issues raised by the experts for discussion by the jury participants, or the participants themselves?</p> <p>P6, line 15 'conference call between expert and jurors for questioning' Clarify the format of this - how were the jurors questions determined ie directly asked of experts by individual participants, asked by the researcher on behalf of questions agreed by the participants among themselves first etc. How did you ensure all juror's questions were addressed, was there any indication that some did not feel confident to participate in the Q&A session with experts?</p> <p>P6, line 31 '3 time points' Why were these time points were selected? I assume the women were not asked their views at the start of the jury process as the researchers were only interested in informed views,.. is this the case?... would be useful to articulate the rationale for the timing decision.</p> <p>P6, line 37 'identify the key reasons' My understanding was this was identified from the final reporting process? Do you mean this analysis was done to further explore and clarify the key reasons presented? It would be useful to report if this analysis revealed something different / additional to what had been explicitly reported by the group in the feedback session. Also to what degree were the participants' rationale and reasoning</p>
-------------------------	---

	<p>about their response to Part A made explicit by them in their reporting, vs what was more implicitly revealed via your analysis of their discourse?</p> <p>P6,line 38 'differences between lay and expert perspectives' Interesting that you identify that the differences between the lay and expert perspectives were key but this is not at all addressed in the results. It left me thinking what were the differences, and in what way were they key? There seems to be a reference to this point in the discussion ie in relation to the clinical literature - is this what you meant? Or different to the experts who presented at the jury. Would be good to clarify and elaborate on this issue in the paper.</p> <p>P7,line 3 – 'cancelled each other out' I read this to mean the overall balance vs the actual number of votes. Also not quite clear what you meant by 'these shifts cancelled each other out'</p> <p>P8,line 43 'extending the organized screening program' The question of further extending the existing organised program to older women was not addressed by the jury (as far as reported here) – or was it? The current program now already offers screening to women aged 70-74.. do you mean to women older than this?</p> <p>P9,line 16 'came to similar conclusions' This would be strengthened if you can also say 'and both juries reported similar reasoning' (or were there any differences in the reasoning between the two juries?)</p> <p>P9,line55 'no matter how well-founded in the evidence' This seems to run counter to your findings that the participants said they would potentially change their mind if the evidence was improved over time, and implications became clearer..</p> <p>P10,line6 'even older women' This sounds a little odd - just not clear what is meant by 'even older women' ... was there prior expectation that older women would have different views to younger women? Or do you mean 'Older women, even those who have been informed in detail...'</p> <p>P10, line10 'Arguments for withdrawal...' Maybe better to clarify you are referring to 'Current arguments' for withdrawal? ie your results report that the women said new or additional evidence may convince them to change their mind.</p>
--	--

REVIEWER	Heleen M.E. van Agt, senior researcher Erasmus Medical Center Rotterdam, Department of Public Health, The Netherlands
REVIEW RETURNED	22-Jan-2018

GENERAL COMMENTS	<p>This is a very interesting paper about the question whether breast cancer screening should be offered to older women despite the mixed evidence about the harms and benefits, was asked to the target groups themselves.</p> <p>Comments: Although explained in the main text, the concepts 'informed views' and 'community jury' may need clarification in the abstract as well. What was the reason for choosing 'whether breast cancer screening should be continue to invite...' as the key issue, as mentioned in the</p>
-------------------------	---

	<p>introduction, rather than ‘whether breast cancer screening should be stop to invite..’? The first option may bring about more arguments relating to the benefits, whereas the second option may create more arguments relating to the harms of breast cancer screening in older women. This should be an issue for the discussion.</p> <p>How was ascertained whether the jury members really understood the harms and benefits of breast screening in older women?</p> <p>p. 5 – participants and recruitment: the majority of members in both juries was from high SES of suburb is an relevant finding, which should be addressed.</p> <p>p. 8 – ‘pre-existing conditions can interact with and compound the harms of standard breast cancer treatment’: what kind of pre-existing conditions are these , can you give an example?</p> <p>p. 8 - Expert’s 4 – What does this mean?</p> <p>Discussion:</p> <p>‘even imperfect information could assist women...’: It is unclear on what results this statement is based.</p> <p>‘Notably, however, this position was amenable to change’....It is unclear whether this statement refers to the current results or to those of another study.</p> <p>The fact that two juries came to the same conclusion and replication of the study will likely produce similar results, does not solve the problem that these views may not be representative: the word ‘however’ does not seem logical.</p> <p>It would be helpful to have an overview of the various reasons to support or reject the presented options, as mentioned in jury 1 and jury 2, for example in Table 3.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Comment and Response BMJ OPEN

	Comment	Response
Editor	Please revise the ‘Strengths and limitations’ section of your manuscript. This section should relate specifically to the methods, and should not include a general summary of, or the results of, the study.	<p>The ‘Strengths and Limitations section of the Discussion has been completely redrafted to relate specifically to the methods used in this study. On P10 Line 9-34 it now reads</p> <p>Possible limitations to this study include 1) the small size of the groups; and 2) the relatively high socio-economic status of the residential areas from which participants came. With respect to group size, however, we note that this is an inevitable characteristic of jury research. Community juries are comprised of small groups of ‘engaged citizens’. Community juries are designed to promote participant inclusivity and deliberative participation rather than achieve statistical representation. Juries are typically comprised of a manageable number of people (12-15) to ensure the quality of participation and deliberation: in larger groups it is difficult to ensure quality of process. The constitution of these juries was in fact a strength. While</p>

		<p>most of the participants lived in areas of relative socio-economic advantage, the rapid gentrification of some areas of Sydney make this socio-demographic distinction difficult to interpret for older age groups. We paid close attention, through screening, to obtaining a socially and culturally diverse sample. Because two juries came to similar conclusions underpinned by similar reasoning, it seems likely our findings are replicable.</p> <p>A possible limitation is the absence of expert testimony from breast cancer patients or survivors. However because all of the expert witnesses have previously occupied or continue to occupy relevant clinical roles, they were able to reflect and comment upon the more personal aspects of breast cancer diagnosis and care in response to jurors' questions and discussions. A strength of this study was the quality and reputation of the experts who gave testimony, and the process by which they moderated one another's presentations until all experts could accept that all views presented could be argued from the evidence.</p>
Rev #1	<p>having a baseline description (if available) of what the public thinks at baseline would present helpful comparison/contrast with the study results.</p>	<p>Our research objective was to capture the views of women in the target age group who have been informed of the potential benefits and harms of screening participation and had the opportunity to discuss the evidence with their peers. So rather than take a baseline before any evidence has been presented or discussed, and potentially explicitly encouraging jurors to take an uninformed position and defend it, we elected to take the first survey at the end of the evidence sessions on Day 1. That said we agree that some indication of public perceptions of mammography screening is warranted so we have introduced the following text to the Methods section on page 6 Line 53-57:</p> <p>Previous research indicates that Australian women are generally enthusiastic about screening but have minimal awareness about overdiagnosis. In our previously reported trial of a decision aid about whether to continue or stop screening among women aged 70 years, 78% reported a positive intention to continue screening at baseline.</p>

2	<p>why two juries (as opposed to three for example) – is it the norm? is it partially because the same verdict was reached in both juries?</p>	<p>The jury is the unit of analysis and therefore having a second jury to act as a comparator allows us to better understand how different values, beliefs, and forms of evidence had influence or otherwise on each group’s deliberations and verdict.</p> <p>This information is not usually included in published reports of CJs – but we are happy to do so if the reviewer or editor believes it be important.</p>
3	<p>mention under limitations that the findings are also influenced by what specific information were presented in the juries, how they were presented, and by which specific experts. For example in retrospect, perhaps a clinician or even a patient instead of an epidemiologist who can comment on more personal aspects of the harms of overdiagnosis and overtreatment may have had more sway with the juries.</p>	<p>The expertise of each expert witness and content of their presentation is displayed in table 2. The presentations shown to the jury are also available online so that readers who are interested can make their own judgement. We appreciate the reviewer’s concerns and designed the study to make sure that clinical perspectives were a key component of the evidence jurors were provided. Consequently, all of the expert witnesses have previously or continue to occupy senior clinical roles in breast cancer screening and treatment. To highlight this we have added the following text to the limitations section in the Discussion on Page 10 Lines 26-33:</p> <p>A possible limitation is the absence of expert testimony from breast cancer patients or survivors. However because all of the expert witnesses have previously occupied or continue to occupy relevant clinical roles, they were able to reflect and comment upon the more personal aspects of breast cancer diagnosis and care in response to jurors’ questions and discussions. A strength of this study was the quality and reputation of the experts who gave testimony, and the process by which they moderated one another’s presentations until all experts could accept that all views presented could be argued from the evidence.</p> <p>The biosketches of the expert witnesses on the project website also make clear their respective clinical credentials – this is made accessible to the reader via reference 24 in the bibliography.</p> <p>As to the possibility of including patient perspectives on</p>

		<p>the harms of screening and overdiagnosis in future studies, as we note on page 9 Line 49-52, one of the conundrums of conducting research in this area is that there is no way to know which individuals are overdiagnosed. Therefore, in designing the CJ, it is not possible to balance the testimony of someone who believes themselves to have benefited from screening by providing equal time to someone who has been harmed. It is also very possible that a presenter who believes herself to have been saved by breast screening could in fact have been overdiagnosed. This makes involvement of consumers as expert witnesses in this area fraught and very difficult to achieve.</p>
4	<p>I wonder if the authors may be able to comment on how the information is presented to the juries may have impacted their opinions. It seems like the juries did not resonate with information that are population-based and value more personal and symbolic meanings around screening. I wonder if, in future studies, presenting for example the harms of screening with more personal stories and examples may lead to different responses.</p>	<p>As noted above, it is difficult to present personal stories of the harms of screening because of the nature of overdiagnosis. Our expert witnesses were able to convey some of the personal aspects in their presentations and during discussions such that jurors were able to consider both population and personal impacts during their deliberations. We agree with the reviewer that most jurors valued more personal and symbolic meanings in their responses – as such the challenge remains communicating about the potential harms of screening such as overdiagnosis without being able to draw on a set of powerful personal counter-narratives.</p>

Rev #2	Comment	Response
1	<p>P4, line 53 ‘they provide evidence’ It would be useful to elaborate / distinguish between epidemiological evidence and the type offered by the findings of community juries – how does it assist policy making</p>	<p>We thank the reviewer for pointing to this ambiguity – we have amended the text on P4 Line 54 to now read:</p> <p>The process is like a legal proceeding, but the outputs are not legally binding: instead they provide evidence of public values and the likely acceptability and perceived legitimacy of different policy alternatives to assist policymaking.</p>
2	<p>P5, line 5 ‘all stake holders agreed’ Why did all stakeholders all agree this was the most important</p>	<p>As we note in the introduction, a policy change was made in 2013 to extend the target age range for screening from 50-69 to 50-74 years. This led to invitations being mailed</p>

	<p>question to address? Was a change of policy being considered? Was a change in current policy being advocated or did some of the groups want to advocate this?</p> <p>Also did the stakeholders and researchers discuss in advance how the findings of this study would be used?</p> <p>It would be useful to include here some more information about the policy context in which this research was being conducted.</p>	<p>to women aged 70-74 nationally, which had not happened previously. Since then there have been debates as to whether evidence of benefit should have been established first, and whether older women would want to participate if they were informed of the risks of overdiagnosis and overtreatment. To ensure readers have an understanding of this policy context and the purpose of our study we have amended the relevant section on Page 5 Lines 3-16 to now read:</p> <p>We consulted with major stakeholders (consumer organisations, epidemiologists, women’s health physicians, and the Cancer Council of Australia) to determine the most appropriate questions for the community juries to consider (Figure 1). Because of continued uncertainty as to the balance of benefits and harms of breast cancer screening in older women, all stakeholders agreed that in the absence of such evidence, the key issue to consider was whether BreastScreen Australia should continue to invite women aged 70-74 to participate in its program. Even though a further change in policy is not currently being considered, the results of this study could be used to inform discussion and provision of information about overdiagnosis and overtreatment of screen detected breast cancer in older women and to inform future research on health communication. Additionally, the results may assist in developing policy in other jurisdictions where changes in the target age group for breast screening are being considered such as the UK.</p>
3	<p>P5, line 51 ‘interrogating the evidence’</p> <p>Would be good to include some more detail of what type of evidence was interrogated (assume this means epidemiological or also other?).</p>	<p>Table 2 describes the content of the expert presentations. In addition to epidemiological evidence, jurors heard evidence on basic cancer biology, and both clinical and practical issues.</p> <p>To make this clearer to the reader we have amended the text in the Methods section on Page 6 Line 12-14 to read:</p> <p>Jury Day 1 focused on interrogating the epidemiological evidence and understanding basic cancer biology and common clinical and practical issues (Table 2).</p>

4	<p>And the 'clinical and practical issues' ... were these issues raised by the experts for discussion by the jury participants, or the participants themselves?</p>	<p>To make clear that both experts and jurors raised issues for discussion during the Q and A sessions we have amended the text on Page 6 Lines 29-33 now read:</p> <p>After each expert's video was screened, we opened a conference call between that expert and the jurors for questioning. Facilitated by a researcher, these question and answer sessions allowed jurors to clarify or challenge the arguments presented, ask further questions, and raise and discuss practical and clinical issues that were important to them.</p>
5	<p>P6, line 15 'conference call between expert and jurors for questioning' Clarify the format of this - how were the jurors questions determined ie directly asked of experts by individual participants, asked by the researcher on behalf of questions agreed by the participants among themselves first etc. How did you ensure all juror's questions were addressed, was there any indication that some did not feel confident to participate in the Q&A session with experts?</p>	<p>We agree that the format and character of the Q and A sessions could be described more clearly. The relevant section on Page 6 Line 33-40 now reads:</p> <p>Participants asked the experts their individual questions directly via telephone, following a brief preparatory discussion among the group. At the end of the interaction with each expert witness, jurors were asked whether they were satisfied or had further questions. Because new issues can emerge and gain importance to jurors during the course of a citizens' jury, expert witnesses remained available via email to answer any further questions that arose during subsequent proceedings. Facilitation focused on promoting constructive dialogue and fair interaction amongst jurors.</p> <p>The results of the Exit Surveys support our impression that all of the jurors were confident in their interactions with each other and the experts. While we avoided formally 'testing' participants so as not to intimidate them, the Exit surveys for both juries indicates that all participants believed that the process was fair and that they had sufficient understanding of the evidence presented to discuss the issues important to them and come to a final decision.</p>
6	<p>P6, line 31 '3 time points' Why were these time points were selected? I assume the women were not asked their views at the start of the jury process as the researchers were only interested in informed views,... is this the case?... would be useful to articulate the</p>	<p>Please see our response to Rev #1 Comment #1 above. To make the rationale for each time point survey clearer we have amended the relevant text on Pages 6 Line 53 to Page 7 Line 8 to read:</p> <p>Previous research indicates that Australian women are generally enthusiastic about screening but have minimal</p>

	<p>rationale for the timing decision.</p>	<p>awareness about overdiagnosis.²⁵ In our previously reported trial of a decision aid about whether to continue or stop screening among women aged 70 years, 78% reported a positive intention to continue screening at baseline.²⁶ To track changes in the positions held by individual jurors, participants completed an anonymous ballot at 3 time-points during jury proceedings (after they had been presented all the evidence at the conclusion of day 1; after they had had time to consider this evidence overnight at the beginning of day 2; and, after the deliberation and delivery of the verdict at the end of day 2). Jurors also completed an Exit Survey for the purposes of process evaluation at the very end of the final jury session.</p>
7	<p>P6, line 37 'identify the key reasons' My understanding was this was identified from the final reporting process? Do you mean this analysis was done to further explore and clarify the key reasons presented? It would be useful to report if this analysis revealed something different / additional to what had been explicitly reported by the group in the feedback session. Also to what degree were the participants' rationale and reasoning about their response to Part A made explicit by them in their reporting, vs what was more implicitly revealed via your analysis of their discourse?</p>	<p>We agree with the reviewer that this sentence is confusing – the key reasons for the verdicts and their reasoning were identified and reported by both deliberative groups during the final session of each jury. The subsequent analysis of jury transcripts was to further explore and clarify the key reasons presented. Our analyses did not reveal any implicit norms or sets of reasons beyond those identified by participants. To make this clear the text on Page 7 Line 10-14 now reads:</p> <p>Jury transcripts were subsequently reviewed to further explore and clarify the key reasons why jurors supported or rejected the presented options. In what follows we have summarized jurors' own descriptions of the rationale and reasoning that underpins their responses to the question asked of them.</p>
8	<p>P6,line 38 'differences between lay and expert perspectives' Interesting that you identify that the differences between the lay and expert perspectives were key but this is not at all addressed in the results. It left me thinking what were the differences, and in what way were they key? There seems to be a reference to this point in the discussion ie in relation to the clinical literature - is this what you meant? Or different to</p>	<p>Our intention was to flag the differences between lay perspectives and those common to epidemiologists (not just those who presented to the jury) on the relative of importance different potential benefits and harms from screening participation – we have changed the text on Page 15 Line 7 to make this clear.</p> <p>Further to the reviewer's comment about making these differences clearer in the results and discussion sections we have added a few sentences to Page 8 Line 9-14</p>

	<p>the experts who presented at the jury. Would be good to clarify and elaborate on this issue in the paper.</p>	<p>For these reasons, jurors argued that decision makers should be cautious about limiting opportunities for early detection. This was because they ascribed a broader set of benefits to screening participation than those commonly recognised by epidemiologists. Invitations to women in this age group, they argued, should cease only when the evidence of an adverse balance of harms to benefits is solid and not contested by experts.</p> <p>And Page 11 Line 16-20</p> <p>Epidemiologically evidenced, population-based information about potential benefits and harms of participation does not appear to resonate sufficiently with many women to lead them reassess the symbolic and personal values and meanings they ascribe to screening.</p>
9	<p>P7,line 3 – ‘cancelled each other out’ I read this to mean the overall balance vs the actual number of votes. Also not quite clear what you meant by ‘these shifts cancelled each other out’</p>	<p>Apologies for the confusion, the concept of “overall balance” is what we are trying to articulate. To make this clearer this section on Page 7 Line 32-36 now reads:</p> <p>Even though the overall balance of votes remained fairly stable during the course of both juries, analysis of the 3 time-point ballots indicates that several participants changed their positions during jury proceedings (3 in July 1 and 5 in July 2). Although some individuals changed their position, the overall majority position of the groups did not change. This is because individuals shifted in both directions – towards and away from supporting screening (Table 3)</p>
10	<p>P8,line 43 ‘extending the organized screening program’ The question of further extending the existing organised program to older women was not addressed by the jury (as far as reported here) – or was it? The current program now already offers screening to women aged 70-74.. do you mean to women older than this?</p>	<p>We thank the reviewer for picking up this confusing sentence – the amended text on Page 9 Lines 36-39 now reads:</p> <p>Some women thus argued that money spent on offering organised screening to women aged 70-74 would be better spent on breast cancer research.</p>

11	<p>P9,line 16 'came to similar conclusions'</p> <p>This would be strengthened if you can also say 'and both juries reported similar reasoning' (or were there any differences in the reasoning between the two juries?)</p>	<p>We once again thank the reviewer for highlighting this lacuna in our description of the results. Both juries did come to similar positions on the basis of similar arguments and reasoning. To reflect this the text on Page 10 Line 22-24 now reads</p> <p>Because two juries came to similar conclusions underpinned by similar reasoning, it seems likely our findings are replicable</p>
12	<p>P9,line55 'no matter how well-founded in the evidence'</p> <p>This seems to run counter to your findings that the participants said they would potentially change their mind if the evidence was improved over time, and implications became clearer..</p>	<p>We agree that this statement is at odds with our findings and have amended the text (to better reflect our central recommendation) on Page 11 Lines 20-23 to now read:</p> <p>Consequently, any changes in the organisation of mammography screening need to be strongly founded in evidence, but are also likely to require greater-than-usual transparency and engagement with other relevant community values.</p>
13	<p>P10,line6 'even older women'</p> <p>This sounds a little odd - just not clear what is meant by 'even older women' ... was there prior expectation that older women would have different views to younger women? Or do you mean 'Older women, even those who have been informed in detail...'</p>	<p>Our intention was to convey the latter – to remedy this we have changed the text on Page 11 Line 28-32 to read:</p> <p>Older women, even those who have been informed in detail of the potential benefits and harms of screening participation, may highly value early detection programs, seeing the invitation to screening as an opportunity for choice and a demonstration that society continues to recognise and invest in them.</p>
14	<p>P10, line10 'Arguments for withdrawal...'</p> <p>Maybe better to clarify you are referring to 'Current arguments' for withdrawal? ie your results report that the women said new or additional evidence may convince them to change their mind.</p>	<p>Thanks for picking this up – as suggested by the reviewer the amended text on Page 11 Line 32-34 now reads:</p> <p>Current arguments for withdrawal of breast screening because of harms associated with overdiagnosis in people with reduced life expectancy seem unlikely to resonate with older women.</p>

Rev #3	Comment	Response
1	Although explained in the main text, the concepts 'informed views' and 'community jury' may need clarification in the abstract as well.	We appreciate that an explanation of the terms 'informed views' and 'community jury' in the abstract would be useful, but the word limit for this section precludes this, and as the reviewer notes both concepts are explained at some length in the main text
2	What was the reason for choosing 'whether breast cancer screening should be continue to invite...' as the key issue, as mentioned in the introduction, rather than 'whether breast cancer screening should be stop to invite..'? The first option may bring about more arguments relating to the benefits, whereas the second option may create more arguments relating to the harms of breast cancer screening in older women. This should be an issue for the discussion.	<p>We agree with the reviewer that there is a subtle difference in asking a CJ whether an activity should <u>continue</u> as opposed to whether it should be <u>stopped</u>, especially if asked of someone who is not informed about the finer details and issues implied by the question. However, because all of the jurors were presented with detailed, balanced and factual evidence, and took part in extended discussions about the potential benefits and harms of breast cancer screening and their implications for continuing or stopping invitations to women aged 70-74 we do not believe that these framing effects had influence on jurors' reasoning or the verdict.</p> <p>Although we cannot provide explicit evidence to this effect, after 2 days of jury proceedings we are confident that all of the jurors understood that continuing to send invitations to their age group meant that the invitations would not stop, and that not continuing to send invitations meant stopping.</p>
3	How was ascertained whether the jury members really understood the harms and benefits of breast screening in older women?	<p>With respect to the reviewer we are struggling to understand what they mean by "really" understood. Because it is not technically possible to identify individuals who have been overdiagnosed and because none of the jurors had experienced breast cancer diagnosis and/or treatment it could be argued that neither they nor anyone in their position can really understand the harms or benefits of screening. In so far as to say that if we applied this standard reflexively, then many experts in the field may also lack this 'real' understanding.</p> <p>If the definition of 'really' understanding the harms and</p>

		<p>benefits of breast screening in older women means that participants comprehended enough information to make a decision that authentically reflects their preferences and values, then we note that the event was organised to inform women through structured provision of evidence and opportunities to discuss and clarify concepts and claims through interactions with several experts. While we avoided formally ‘testing’ participants so as not to intimidate them, the Exit survey for both juries indicates that all participants believed that the process was fair and they were sufficiently informed to make a decision.</p> <p>So as to make this clearer to the reader we have added the following text to Page 7 Line 36-45</p> <p>Transcripts of questions and discussions during proceedings indicate that over the course of the jury proceedings the vast majority of the jurors comprehended concepts being discussed and that all of them understood the trade-offs implicit in the question we were asking them to address during their deliberations. While we avoided formally ‘testing’ participants so as not to intimidate them, the Exit survey for both juries show that all participants believed that the process was fair and that they had sufficient understanding of the evidence presented to discuss the issues important to them and come to a final decision.</p>
4	<p>p. 5 – participants and recruitment: the majority of members in both juries was from high SES of suburb is an relevant finding, which should be addressed.</p>	<p>We acknowledge that the higher SES of most participants should be commented on in the main text in the Methods section on Page 5 Line 52- Page 6 Line 7 which now reads</p> <p>We contracted an independent professional research service to recruit two juries of women aged 70-74 living in Greater Sydney, Australia from randomly generated list-based samples and random digit dialing. We selected women based on their socio-demographic characteristics, as well as their eligibility and availability. Because women born in the late 1940s are likely to have had more limited educational opportunities than subsequent generations, and because many of them are long-term residents of recently gentrified suburbs, we chose to prioritise the stratification of participant education levels in our</p>

		<p>recruitment strategy. Potential participants with a personal history of breast cancer (themselves or close family member) were excluded through a screening interview, as were health professionals and those working in breast cancer advocacy. 34 women were recruited (Table 1). The juries were socially and culturally diverse, sampling was skewed towards higher levels of socioeconomic advantage and lower educational attainment than the average for the Australian population (Table 1). All jurors received a modest honorarium in recognition of their participation and contribution to jury processes and outcomes.</p> <p>We have also noted this in the discussion of limitations on Page 10 Line 17-24 where it now reads</p> <p>While most of the participants lived in areas of relative socio-economic advantage, the rapid gentrification of some areas of Sydney make this socio-demographic distinction difficult to interpret for older age groups. We paid close attention, through screening, to obtaining a socially and culturally diverse sample.</p>
5	p. 8 – ‘pre-existing conditions can interact with and compound the harms of standard breast cancer treatment’: what kind of pre-existing conditions are these , can you give an example?	<p>We thank the reviewer for pointing out this opportunity to provide more information. We have amended the text on Page 9 Line 19-23 to now read</p> <p>Participants who took this position in both juries gave great significance to evidence that pre-existing conditions such as heart disease and preclinical cognitive disorders (which may be unknown to the individual affected) can interact with and compound the harms of standard breast cancer treatments.</p>
6	p. 8 - Expert’s 4 – What does this mean?	<p>We agree that our expression is clunky. The new text on Page 9 line 38-40 now reads</p> <p>These women also tended to endorse the proposal put forward by Expert 4 that clinical examination was a more trustworthy means of detection in older women.</p>
7	‘even imperfect information could assist women...’: It is unclear on	<p>We thank the reviewer for picking this up ... the text on</p>

	what results this statement is based.	<p>Page 9 Line 54-56 now reads:</p> <p>In their deliberations several jurors argued that even imperfect information could assist women to make their own choices.</p>
8	Notably, however, this position was amenable to change'....It is unclear whether this statement refers to the current results or to those of another study.	<p>Thanks again for highlighting this ambiguity – new text Page 9 line 55 – Page 10 Line 5 reads</p> <p>Notably, however, during the reporting of the verdicts jurors also sought to emphasise that their support for this position was amenable to change. Many jurors who voted to continue to invite women now said if the current UK age extension trial found definitive evidence of significant harms from screening participation they might alter their position.⁷</p>
9	The fact that two juries came to the same conclusion and replication of the study will likely produce similar results, does not solve the problem that these views may not be representative: the word 'however' does not seem logical.	<p>We do not claim that the views of the two juries are representative. To reinforce this point we have added the following text to Page 10 Line 10-24</p> <p>With respect to group size, however, we note that this is an inevitable characteristic of jury research. Community juries are comprised of small groups of 'engaged citizens'. Community juries are designed to promote participant inclusivity and deliberative participation rather than achieve statistical representation. Juries are typically comprised of a manageable number of people (12-15) to ensure the quality of participation and deliberation: in larger groups it is difficult to ensure quality of process. The constitution of these juries was in fact a strength. While most of the participants lived in areas of relative socio-economic advantage, the rapid gentrification of some areas of Sydney make this socio-demographic distinction difficult to interpret for older age groups. We paid close attention, through screening, to obtaining a socially and culturally diverse sample. Because two juries came to similar conclusions underpinned by similar reasoning, it seems likely our findings are replicable.</p>
10	It would be helpful to have an overview of the various reasons to	We appreciate the reviewer's desire for accessible data representation as to the reasons each jury held for and

	support or reject the presented options, as mentioned in jury 1 and jury 2, for example in Table 3.	against the presented options. We have included a new Table (#4) which contains examples of the reasons participants gave for and against continuing to invite women aged 70-74 to participate in mammography screening
--	---	---

VERSION 2 – REVIEW

REVIEWER	Lucie Rychetnik School of Medicine Sydney, University of Notre Dame Australia
REVIEW RETURNED	21-Mar-2018

GENERAL COMMENTS	The authors have done an excellent job at addressing the reviewers' comments and I am happy to recommend publication.
-------------------------	---

REVIEWER	Heleen M.E. van Agt, senior researcher Erasmus MC University Medical Center, Department of Public Health, Rotterdam, the Netherlands
REVIEW RETURNED	10-Apr-2018

GENERAL COMMENTS	I agree with the author's responses and revisions in the manuscript.
-------------------------	--

REVIEWER	Nancy Schoenborn Johns Hopkins University, USA
REVIEW RETURNED	12-Apr-2018

GENERAL COMMENTS	The authors have adequately addressed my prior comments. I do not have any further comments.
-------------------------	--