

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([see an example](#)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	A Community Jury on PSA Screening: what do well-informed men want the government to do about prostate cancer screening- a qualitative analysis
<b>AUTHORS</b>	Rychetnik, Lucie; Doust, Jenny; Thomas, Rae; Gardiner, Robert; MacKenzie, Geraldine; Glasziou, Paul

### VERSION 1 - REVIEW

<b>REVIEWER</b>	A/Prof Jennifer Whitty University of Queensland, Australia
<b>REVIEW RETURNED</b>	28-Feb-2014

<b>GENERAL COMMENTS</b>	<p>This is well written and interesting paper presenting findings from a community jury undertaken in Australia, which aimed to determine the priorities, values and concerns of men aged 50-70 years around the benefits and harms of PSA screening, as well as to examine the value of the community jury process in eliciting public values on PSA screening. The paper presents consumer focussed evidence on a controversial topic of high relevance to community practice.</p> <p>I have several comments and suggestions on which the authors might like to reflect.</p> <ul style="list-style-type: none"><li>• I would suggest the authors consider adding a brief introduction to the community jury as a methodological approach to the start of the methodological section, along with a justification for choice of this deliberative method over others. This would aid readers who are not familiar with this and other deliberative engagement methods.</li><li>• Why was there uneven randomisation (12 participants randomised to the jury and 15 to the control group)? I would think with the randomisation method (paper draw) a more even randomisation could have been achieved.</li><li>• Could the authors expand on and justify their approach to statistical analysis for the survey data? For example, the pre and post survey are repeated measures in the same individuals – was repeated measure ANOVA used? Similarly, the regression analysis would need to consider the potential for correlation in the repeated measures.</li><li>• It strikes me that the (self-selecting) men participating may have been particularly anxious or have particular preconceptions around PSA screening that might not be typical of men of their age range in Australia. Could the authors comment on how typical the men participating of men who might consider screening are? For example, we are told the results of prior testing – what is the current uptake of PSA screening in men of this age group in Australia? This is relevant for considering the generalizability of results.</li><li>• The discussion is quite Australia centric. Could the authors place their findings in the context of what is known of mens' perceptions</li></ul>
-------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

	<p>and opinions around screening internationally? Are the opinions of men in Australia likely to be relevant outside Australia, or what messages can an international audience gather from the study?</p> <ul style="list-style-type: none"> <li>• A member of the research team and co-author (PG) was also an expert witness for the jury (against screening). The jury as a group seem to have moved their position to a more negative opinion of screening over the course of the jury. Clearly witnesses have the potential to influence this (as would be expected given the method). However, I wonder whether there was any potential for witnesses to have unequal influence on the jurors. For example, was PG present during any part of the jury other than the expert presentation? Was it declared to participants that PG was a member of the team? If either of these is the case, there would seem to be potential for the jury to respond in a manner akin to social desirability bias, and “please” the researchers. I would suggest the authors reflect on whether the potential influence of the research team could be worthy of mention as a potential limitation of their study.</li> </ul> <p>Thank you for the opportunity to review this paper.</p>
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<b>REVIEWER</b>	Michael Burgess University of British Columbia, Canada
<b>REVIEW RETURNED</b>	05-Mar-2014

<b>GENERAL COMMENTS</b>	<p>this is generally a well-written paper about a well-conceived project. My suggestions for improvement are modest. Some of my comments below explain the ratings above and do not require revisions.</p> <p>The methods are not described or referenced in detail, but they are widely understood and available in the literature.</p> <p>There is a vast theoretical and methodological literature on deliberative engagement and citizen juries, but it is probably beyond the scope of the is manuscript to cite.</p> <p>The results would benefit from being distinguished in terms of which are based on the qualitative findings and which are the jury conclusions.</p> <p>The study limitations are barely mentioned in passing, although space limitations may be a factor</p> <p>I offer these comments to suggest ways pin which the authors might respond to some of my concerns above, and for general consideration. I suggest few specific revisions because they may not be appropriate for this article due to space and audience.</p> <p>Page 6. Lines 10-30: Recruitment is in general inevitably biased by self-selection and opportunity, but the study design does nothing to mitigate these biases, and introduces unnecessary ones. While the age of the men recruited to the exercise reflect the at-risk age, it is not obvious why men from other age groups would be excluded. Seeking deliberative input from other age ranges may reveal different risk tolerances that could be relevant to the policy recommendations. Recruitment into a participant group that was stratified by demographic variables or initial questionnaires to ensure diverse life perspectives would have increased the extensibility of</p>
-------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

the study across diverse populations, and might have enhanced the diversity of interests considered in the Jury. On the other hand, attracting a more disinterested participant group might have required payment to motivate participation, increasing the cost of the study, and perhaps requiring assistance with recruiting.

Given that the paper later discusses the role of wives and families, it might have been interesting to include some in the discussion.

There is very brief mention of varying the method of recruitment on page 19, but without any analysis of how it might have affected the deliberation process or conclusions.

Page 6, line 59 in Box 1: It is not clear from the paper what “a government organized invitation program for testing for prostate cancer” means. How was this introduced and explained? Is it the invitation or the testing program that is government organized (is it advertising to seek testing through GPs or an offer of direct access to testing that circumvents the GP)?

Page 7, What is the effect of the very expert and cognitively driven information sessions? There are significant critiques of deliberative methods for failing to present and legitimate other forms of understanding and warranting information. For example, hearing about experiences about living with risk might lead to a different weighting on the value of PSA testing, or give the jury members an opportunity to reflect on whether such experiences should make a difference. This is briefly mentioned on page 19, but without any reflection on the effect it might have had on the deliberation and conclusions.

Page 8: It seems like there are two sets of findings: the jury recommendations and the qualitative findings. It would be good to make this distinction clear, and consider how the two different findings relate to each other. What if the qualitative findings are inconsistent with a jury conclusion or suggest that a recommendation from the jury was based on a crucial misunderstanding, or heavily influenced by one individual?

It also seems like the reader should not have to try so hard to understand when the discussion is based on qualitative analysis and when it is based on the jury conclusions. The qualitative findings are presumably the basis for the discussion on page 10 (reasons for attending the jury), page 11 (core values related to PSA screening), page 12 (the discussion and Box 2 on surprising information). I think the discussion on pages 13-17 shifts to the jury conclusions, including Boxes 3 and 4. It would be helpful to make it clear when the qualitative analysis is the basis for the discussion and when the conclusions of the jury are being discussed.

Page 9, lines 52-58: How does the experience with blood tests for prostate cancer of the men on the jury compare to the population? Was this a typical group, or did they reflect a bias in relation to likelihood to support/resist testing?

Page 12, line 30. others should be others’

Additional points could be made about the ethical dimensions of the deliberation. While the men focused on individual choice, they also surprised the researchers by focusing on the need to inform GPs as

	<p>trusted sources of information and advice. As the authors point out, the jury participants considered higher-order questions about the common good.,although the manuscript does not point out that this is particularly interesting since all the participants had themselves been tested and had clear interests in the decisions they were considering.</p> <p>The authors suggest that it is “important however to repeat the process with other juries to assess the repeatability and generalizability of findings to other parts of Australia.” This is one of the most common critiques of small-group deliberative democracy—that it is too small a group to represent the population. The easiest way to respond is to acknowledge the limitation.</p> <p>But this misunderstands the justification for small group deliberative processes. Representative groups, if by representative is meant a random sample that is sufficiently large to achieve statistical representation of the population, are difficult if not impossible to inform and lead through a deliberative process. As Goodin and Dryzek suggest in their seminal paper on mini-publics, the small group deliberation approximates what a public would conclude, if they were informed and deliberative. It is on the basis of being informed and deliberative that the conclusions have some legitimacy as a policy input. Each new deliberation is more input, but deliberations are not repeatable or generalizable by their design and justification. It might be possible to seek wider public input on the deliberative conclusions or process, and whether the wider public trusts both. Fishkin’s deliberative polling seeks to use a statistical sample of the population. But even this would pose a problem for this research because it is not clear how to define the population relevant to the issue of PSA testing. Should it be the men at risk, all males, or the entire population who form the social context and source of public funding and opportunity costs?</p>
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## VERSION 1 – AUTHOR RESPONSE

### Reviewer 1

*This is well written and interesting paper presenting findings from a community jury undertaken in Australia, which aimed to determine the priorities, values and concerns of men aged 50-70 years around the benefits and harms of PSA screening, as well as to examine the value of the community jury process in eliciting public values on PSA screening. The paper presents consumer focussed evidence on a controversial topic of high relevance to community practice.*

*I have several comments and suggestions on which the authors might like to reflect.*

*I would suggest the authors consider adding a brief introduction to the community jury as a methodological approach to the start of the methodological section, along with a justification for choice of this deliberative method over others. This would aid readers who are not familiar with this and other deliberative engagement methods.*

Response 1: We have added new text to introduce the methods (p6), and included a new reference (#21):

*Why was there uneven randomisation (12 participants randomised to the jury and 15 to the control group)? I would think with the randomisation method (paper draw) a more even randomisation could have been achieved.*

Response 2: We sought to recruit 12 jury participants (ideal jury size). Of 31 men initially recruited 4 withdrew prior to the information evening / group allocation. As all 27 of those who attended the information evening chose to participate in the study randomisation, we placed 12 jury group slips and 15 control group slips in the bucket. The subsequent drop-out from the jury of 2 men (between the time of group allocation on Fri and the final verdict on Sunday) was not anticipated, and in hindsight it may have been better to allocate 14 to the jury.

We have added new text (p6) to clarify this process.

*Could the authors expand on and justify their approach to statistical analysis for the survey data? For example, the pre and post survey are repeated measures in the same individuals – was repeated measure ANOVA used? Similarly, the regression analysis would need to consider the potential for correlation in the repeated measures.*

Response 3: We have inserted text to clarify that a repeated measures ANOVA was used (P10)

We conducted a regression analyses as this is considered the most appropriate statistical tool for controlled trials with pre and post-test measures (Vickers AJ, & Altman DG. Statistics Notes: Analysing controlled trials with baseline and follow-up measurements. (*BMJ* 2001;323:1123-1124). This analysis accounts for correlations by partially adjusting for baseline measurement. We have added this reference to the methods on p10 (ref #28).

*It strikes me that the (self-selecting) men participating may have been particularly anxious or have particular preconceptions around PSA screening that might not be typical of men of their age range in Australia. Could the authors comment on how typical the men participating of men who might consider screening are? For example, we are told the results of prior testing – what is the current uptake of PSA screening in men of this age group in Australia? This is relevant for considering the generalizability of results.*

Response 4: We have revised the discussion (p18-19) to compare the high rate of prior testing among the jury to Australian rates more broadly, and to consider the implications. We have explicitly discussed that juries are intended to reveal the views of different types of 'mini-public', and that while it is not the purpose of a jury to be representative in a statistical or electoral sense, the views of other types of public may also need to be considered. This is to also address the aligned points raised by reviewer 2.

*The discussion is quite Australia centric. Could the authors place their findings in the context of what is known of mens' perceptions and opinions around screening internationally? Are the opinions of men in Australia likely to be relevant outside Australia, or what messages can an international audience gather from the study?*

Response 5: We have added to the expanded discussion (p19) reference to PSA testing rates in both Australia and the US, and noted the need to examine these issues with other publics – both in Australia and other countries with different information and services.

*A member of the research team and co-author (PG) was also an expert witness for the jury (against screening). The jury as a group seem to have moved their position to a more negative opinion of screening over the course of the jury. Clearly witnesses have the potential to influence this (as would be expected given the method). However, I wonder whether there was any potential for witnesses to have unequal influence on the jurors. For example, was PG present during any part of the jury other than the expert presentation? Was it declared to participants that PG was a member of the team? If either of these is the case, there would seem to be potential for the jury to respond in a manner akin to social desirability bias, and "please" the researchers. I would suggest the authors reflect on whether the potential influence of the research team could be worthy of mention as a potential limitation of their study.*

Response 6: The expert witnesses who presented both for and against screening (PG and RG) are co-authors on the paper. They were introduced to the jury as members of the research team, but were

present only for their own talk and Q&A session, and did not attend any of the subsequent jury deliberations.

We have edited the manuscript to clarify these points by explicitly identifying PG and RG as authors and adding new text to first paragraph on p8.

## **Reviewer 2**

*This is generally a well-written paper about a well-conceived project. My suggestions for improvement are modest. Some of my comments below explain the ratings above and do not require revisions.*

*The methods are not described or referenced in detail, but they are widely understood and available in the literature.*

*There is a vast theoretical and methodological literature on deliberative engagement and citizen juries, but it is probably beyond the scope of this manuscript to cite.*

**Response:** We agree that it is beyond the scope of this manuscript to include a large number of additional references on the theoretical and methodological literature on deliberative methods, but we have added a brief introduction to the methods (p6), and the reference for Goodin and Dryzen (#21) as we agree with Reviewer 2 that this is a particularly relevant and helpful reference on the topic of community juries.

*The results would benefit from being distinguished in terms of which are based on the qualitative findings and which are the jury conclusions. (addressed below)*

*The study limitations are barely mentioned in passing, although space limitations may be a factor (addressed below)*

*I offer these comments to suggest ways in which the authors might respond to some of my concerns above, and for general consideration. I suggest few specific revisions because they may not be appropriate for this article due to space and audience.*

*Page 6. Lines 10-30: Recruitment is in general inevitably biased by self-selection and opportunity, but the study design does nothing to mitigate these biases, and introduces unnecessary ones. While the age of the men recruited to the exercise reflect the at-risk age, it is not obvious why men from other age groups would be excluded. Seeking deliberative input from other age ranges may reveal different risk tolerances that could be relevant to the policy recommendations. Recruitment into a participant group that was stratified by demographic variables or initial questionnaires to ensure diverse life perspectives would have increased the extensibility of the study across diverse populations, and might have enhanced the diversity of interests considered in the Jury. On the other hand, attracting a more disinterested participant group might have required payment to motivate participation, increasing the cost of the study, and perhaps requiring assistance with recruiting.*

**Response:** We agree there are both strengths and limitations in the recruitment strategy adopted and the available options. We have expanded the discussion (p18-19) to highlight that this jury was a self selected group of men interested in PSA testing, and to include the other options that may be considered.

*Given that the paper later discusses the role of wives and families, it might have been interesting to include some in the discussion.*

Response: We have added reference to the important role of wives and families to the discussion (p19). We have noted the men had highlighted the influence of partners on their health seeking behaviours, and that this offers a rationale for considering them as an important public to engage in future jury options.

*There is very brief mention of varying the method of recruitment on page 19, but without any analysis of how it might have affected the deliberation process or conclusions.*

Response: We have expanded the discussion to include examples of the alternative methods of recruitment (p19). We have noted that these variations may impact on the deliberations or conclusions of the jury, but prefer not to speculate here on the nature of these impacts, but rather to suggest that this should be empirically studied.

*Page 6, line 59 in Box 1: It is not clear from the paper what “a government organized invitation program for testing for prostate cancer” means. How was this introduced and explained? Is it the invitation or the testing program that is government organized (is it advertising to seek testing through GPs or an offer of direct access to testing that circumvents the GP)?*

Response: The community-level questions presented to the jury for deliberation were purposefully broad and the men were invited to consider what the various options may comprise. The organised invitation program for testing for prostate cancer was introduced as a comparison to mammography screening for women e.g. ‘such as when women are invited to participate in screening for breast cancer’. The men appeared to understand the ‘equity’ rationale for including this consideration. We have revised the text on p6-7 to make this clarification.

*Page 7, What is the effect of the very expert and cognitively driven information sessions? There are significant critiques of deliberative methods for failing to present and legitimate other forms of understanding and warranting information. For example, hearing about experiences about living with risk might lead to a different weighting on the value of PSA testing , or give the jury members an opportunity to reflect on whether such experiences should make a difference. This is briefly mentioned on page 19, but without any reflection on the effect it might have had on the deliberation and conclusions.*

Response: We have added to the expanded discussion the potential impact of presenting other forms of evidence (p19). As for different methods of recruitment, we have noted that this may impact on jury deliberations or conclusions, but prefer not to speculate here on the nature of the impact but rather to suggest that this should be empirically studied.

*Page 8: It seems like there are two sets of findings: the jury recommendations and the qualitative findings. It would be good to make this distinction clear, and consider how the two different finding relate to each other. What if the qualitative findings are inconsistent with a jury conclusion or suggest that a recommendation from the jury was based on a crucial misunderstanding, or heavily influenced by one individual?*

*It also seems like the reader should not have to try so hard to understand when the discussion is based on qualitative analysis and when it is based on the jury conclusions. The qualitative findings are presumably the basis for the discussion on page 10 (reasons for attending the jury), page 11 (core values related to PSA screening), page 12 (the discussion and Box 2 on surprising information). I think the discussion on pages 13-17 shifts to the jury conclusions, including Boxes 3 and 4. It would be helpful to make it clear when the qualitative analysis is the basis for the discussion and when the conclusions of the jury are being discussed.*

Response: We have added new text (p9) to clarify which sections of the results refer to the qualitative findings and which section is about the jury's conclusions.

*Page 9, lines 52-58: How does the experience with blood tests for prostate cancer of the men on the jury compare to the population? Was this a typical group, or did they reflect a bias in relation to likelihood to support/resist testing?*

Response: This has been addressed in our revised text in the discussion (p18-19) – as indicated in our response #6 to reviewer 1.

*Page 12, line 30. others should be others'*

Response: Corrected

*Additional points could be made about the ethical dimensions of the deliberation. While the men focused on individual choice, they also surprised the researchers by focusing on the need to inform GPs as trusted sources of information and advice. As the authors point out, the jury participants considered higher-order questions about the common good, although the manuscript does not point out that this is particularly interesting since all the participants had themselves been tested and had clear interests in the decisions they were considering.*

Response: We have revised the discussion (p18-19) to more clearly identify that this jury was a group of men who were particularly interested in PSA testing and the majority of whom had been previously tested.

We also highlight that this is particularly interesting in relation to their capacity to consider higher order questions about the common good.

*The authors suggest that it is “important however to repeat the process with other juries to assess the repeatability and generalizability of findings to other parts of Australia.” This is one of the most common critiques of small-group deliberative democracy—that it is too small a group to represent the population. The easiest way to respond is to acknowledge the limitation.*

*But this misunderstands the justification for small group deliberative processes. Representative groups, if by representative is meant a random sample that is sufficiently large to achieve statistical representation of the population, are difficult if not impossible to inform and lead through a deliberative process. As Goodin and Dryzek suggest in their seminal paper on mini-publics, the small group deliberation approximates what a public would conclude, if they were informed and deliberative. It is on the basis of being informed and deliberative that the conclusions have some legitimacy as a policy input. Each new deliberation is more input, but deliberations are not repeatable or generalizable by their design and justification. It might be possible to seek wider public input on the deliberative conclusions or process, and whether the wider public trusts both. Fishkin’s deliberative polling seeks to use a statistical sample of the population. But even this would pose a problem for this research because it is not clear how to define the population relevant to the issue of PSA testing. Should it be the men at risk, all males, or the entire population who form the social context and source of public funding and opportunity costs?*

Response: As indicated above and in response to Reviewer 1, we have revised the discussion on p19-20 to more clearly address the issues of generalizability of community jury findings.

We have also added reference to Goodin and Dryzen (#21) to the methods (p6) and incorporated more clearly into our discussion their articulation of the role of community juries as ‘mini-publics’. We agree that to use the term ‘repeatability’ can be misleading in relation to the outputs of a community jury, and have deleted this term in the discussion (p19).



