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)	Doctors' perspectives on PSA testing illuminate established differences in prostate cancer screening rates between Australia and the United Kingdom: A qualitative study
AUTHORS	Pickles, Kristen; Carter, Stacy; Rychetnik, Lucie; Entwistle, Vikki

VERSION 1 - REVIEW

REVIEWER	Sam Farah Royal Australasian College of Surgeons Australia
REVIEW RETURNED	19-Apr-2016

GENERAL COMMENTS	It would have perhaps been better for the authors to develop a standard survey to deliver to participants following focus group discussion, rather than the current study design. There should be some more discussion in regards to various recommendations from specialists college's such as the Urological Society of Australia and New Zealand (USANZ). I don't feel the concerns about selection bias can be addressed adequately with the current sample size.

REVIEWER	MG Kirby University of Hertfordshire and The Prostate Centre London Member of the Prostate Cancer Risk Management Programme advisory group UK
REVIEW RETURNED	10-May-2016

GENERAL COMMENTS	This is a well written comprehensive review of PSA testing (rather than screening) in two countries. The UK data is somewhat out of date and the relevant papers are attached. The discussion would be enhanced by reference to the up to date data. https://www.gov.uk/guidance/prostate-cancer-risk-management-programme-overview PSA cut off levels should be discussed.(they have changed from age related in the UK to a standard cut off for all ages). The reviewer also provided a file in addition to these comments. Please contact the publisher for full details.

REVIEWER	Roman Gulati Fred Hutchinson Cancer Research Center, USA

REVIEW RETURNED	16-Jun-2016

GENERAL COMMENTS	This study presents results of semi-structured interviews conducted with GPs in Australia and the UK who were prospectively recruited to report on practices and perspectives around PSA screening for prostate cancer in asymptomatic men. The comparison between these two settings is interesting, and the paper is clearly written. However, the value of the study as a scientific contribution is very limited. Essentially, the authors informally explored comments by a small number of GPs in either setting and wove these into a story about factors that drive differential practice patterns in general. While the narrative appeared to be plausible, generalizations to national practice patterns and implied causal relationships are not supported by their study design or data. 1. Can the authors estimate how many GPs were contacted to participate in the study? What was the response rate? Given that the total sample sizes were 40 in Australia and 29 in the UK, I would guess that the response rate is very low. Consequently, it is very difficult to expect the data to be representative of general practice. 2. What is meant by the statement (p. 8) "the schedule was modified between interviews based on the developing analysis"? Does this mean the questions that were asked, the order of the questions, the time of day or day of the week, or something else was changed during the course of the study? At a minimum, the bias introduced by this variability should be acknowledged. As a related point, the authors report nearly 4-fold variation in interview durations. Again, it is difficult to know what impact these differences might have on the comments received. 3. All interviews in Australia were completed months before interviews in the UK started. Even if there were few high-profile publications or news reports about PSA screening between the interviews in the two settings, it is impossible to know the effects of collecting information at different time points. As a consequence of the multiple biases, the story told about the main differences in the two settings and what should be changed to impact PSA screening practices is completely speculative. The authors acknowledge that "some selection bias is possible", but this is a gross understatement. The magnitude of selection bias cannot be known and in fact may be severe. Other biases, mostly related to poor study design (e.g., non-standardized interviews and asynchronous collection), should not be disregarded. This is not to say the study is useless. But the authors should recognize the study for what it is: a pilot study that provides useful information about recruiting participants, scheduling interviews, questions to ask, etc. that can be used to conduct a well-designed study to facilitate reliable comparisons between the two settings. I strongly encourage the authors to recognize the biases and limitations of their study and to overhaul the text accordingly. Discussion about what are the key drivers of practice patterns and physician perspectives, recommendations for policy, etc. should all be removed, or possibly relegated to appropriately qualified comments in the Discussion.

	Minor comments 1. Comparisons of incidence and mortality rates between the two settings leave out any differences in primary treatment and follow-up care, but these must be mentioned even if only to observe that they are similar (if this is the case). 2. "Grounded theory", which may be unfamiliar to many readers, is mentioned at the start of the methods with no explanation. If this is retained at all, I suggest moving it to the Discussion where it should be clearly explained and the speculative nature of any commentary related to the small selective samples in this study should be explicitly stated. 3. What exactly is meant by symptomatic vs asymptomatic men? For example, a man with a first-degree family relative with prostate cancer may be more likely to request a PSA test and to receive it in both settings, yet in both settings he would be considered "asymptomatic", correct?

VERSION 1 – AUTHOR RESPONSE

REVIEWER #1:

It would have perhaps been better for the authors to develop a standard survey to deliver to participants following focus group discussion, rather than the current study design.

The research questions for this project were suitable for a qualitative study. This is because:
1. We wanted to understand how GPs explain how(if) they use PSA testing in their practice, and why;
2. We were interested in context – the conditions under which PSA testing does or does not occur in general practice; and
3. We sought to identify drivers of variation, from the GP perspective.
The work that has been done in this area has largely been quantitative surveys and analysis/audit of medical records. While they have been able to show varied rates of PSA testing, GP knowledge and attitudinal factors associated with more or less testing, and the influence of various demographic factors, they have left several questions unanswered, particularly in relation to GPs' reasoning for how they use (or do not use) PSA testing in their practice, and what it is that shapes their approach to screening.
We intend to follow up our qualitative findings with some quantitative work now that we have a nuanced understanding of what is happening in practice and the particular issues relevant to GPs.

There should be some more discussion in regards to various recommendations from specialists College's such as the Urological Society of Australia and New Zealand (USANZ).
This is an excellent suggestion; we have added the USANZ and EUA recommendations to Table 1.

I don't feel the concerns about selection bias can be addressed adequately with the current sample size.
This was a qualitative study. The qualitative methodology we used has allowed us to understand variation in GP's practices, and the reasons they have for those practices, in depth. What our qualitative methodology does not allow us to do is to claim statistical representativeness to a particular population, as the participants were selected for their ability to be informative, rather than because they were statistically representative. Qualitative sampling continues until the practices in question are well understood, which generally requires recruiting different kinds of participants to

allow comparison between participants. In this study, recruitment was shaped to include GPs from the same practice or locality, interstate GPs, GPs from rural, regional, and remote locations, and lastly GPs across England and Scotland, to explore prostate screening practice in a jurisdiction with comparatively lower rates of screening than Australia. This is standard practice in qualitative research, as is acknowledged in all qualitative research text [see (1-4)]

Our sample size is in fact large for a qualitative study, because we continued to recruit to allow richer comparisons to be made. For example, Creswell 1998 advises a minimum of 20-30 (5); Morse 1994 advises 30-50 interviews (6); Charmaz 25 (7); and Green & Thorogood 20 (8). These are all standard and well respected texts, our larger sample size (n=69) demonstrates the rigour with which we continued to pursue comparisons until theoretical saturation was reached.

With respect to the potential for selection bias, our data demonstrates that this is highly unlikely. As shown in our analysis, we heard a very wide and conflicting range of views, expressing very different perspectives on PSA testing. This diversity suggests that it is very unlikely that our sample was biased towards a particular explanation or view of PSA testing.

REVIEWER #2:
The UK data is somewhat out of date and the relevant papers are attached. The discussion would be enhanced by reference to the up to date data.
We thank the reviewer for this comment and for taking the time to attach the relevant references.
We have included the following information in the introduction.
Clinical practice in the UK and Australia is grounded in the same evidence base and international literature, yet the two jurisdictions have notably different rates of PSA testing. In the UK, a study of six English cities reported the annual practice-based PSA testing rate for 2007 (in men aged 45–89 years) to be 6.2% (9). A more recent study analysed data from patient (men aged 45-84 years) electronic records in primary care. It reported for every 100 men enrolled with a GP for one year, 5.03 (asymptomatic men) were tested in 2010, and the rate increased by 8% in 2011 to 5.45 per 100 (10). Note that the data this analysis was based on represents only 5% of the population in England and may not be representative of all practices.

PSA cut off levels should be discussed (they have changed from age related in the UK to a standard cut off for all ages).
We recognise that the revised recommendation regarding referral pathways to specialist services in the UK is an important policy issue, however we consider it beyond the scope of this paper, which is focused on testing practices rather than referral practices.

REVIEWER #3:

Can the authors estimate how many GPs were contacted to participate in the study? What was the response rate? Given that the total sample sizes were 40 in Australia and 29 in the UK, I would guess that the response rate is very low. Consequently, it is very difficult to expect the data to be representative of general practice.
As noted in response to Reviewer #2, our sample was large for a qualitative study, and followed standard qualitative sampling methodology.

What is meant by the statement (p. 8) "the schedule was modified between interviews based on the developing analysis"?
Does this mean the questions that were asked, the order of the questions, the time of day or day of the week, or something else was changed during the course of the study? At a minimum, the bias introduced by this variability should be acknowledged.
As a related point, the authors report nearly 4-fold variation in interview durations. Again, it is difficult

to know what impact these differences might have on the comments received.

Asynchronous data collection and data analysis is a key feature of a grounded theory methodology and is a strength rather than a weakness. Each participant is selected because of the particular information they can contribute. The interview is responsive to their expertise and ability to inform the analysis. This provides much higher quality, more insightful data, as the interviewer can continue questioning until each point made by each interviewee is fully explained and understood. In contrast, standardised questioning tends to produce thin answers in which much is assumed or unclear. Responsive and flexible data collection is the methodological standard in qualitative inquiry (1-4).

All interviews in Australia were completed months before interviews in the UK started. Even if there were few high-profile publications or news reports about PSA screening between the interviews in the two settings, it is impossible to know the effects of collecting information at different time points.As shown in our analysis, our focus was on discovering the main considerations and concerns of the participants, and the reasoning behind their practice. If we had been quantitatively measuring and comparing, for example, the numerical PSA result at which GPs referred to urology, then it would have been important to measure everyone at the same time. However, as this was not the type of research question we were asking, the short time period between Australian and UK interviews is unlikely to have had any impact.

As a consequence of the multiple biases, the story told about the main differences in the two settings and what should be changed to impact PSA screening practices is completely speculative. The authors acknowledge that "some selection bias is possible", but this is a gross understatement. The magnitude of selection bias cannot be known and in fact may be severe. Other biases, mostly related to poor study design (e.g., non-standardized interviews and asynchronous collection), should not be disregarded.

Please refer to response to reviewer 1 re. selection bias and justification for our qualitative study design, which, as explained, is standard practice in qualitative health research around the world, including research published in this journal and in the BMJ.

The authors should recognize the study for what it is: a pilot study that provides useful information about recruiting participants, scheduling interviews, questions to ask, etc. that can be used to conduct a well-designed study to facilitate reliable comparisons between the two settings. I strongly encourage the authors to recognize the biases and limitations of their study and to overhaul the text accordingly. Discussion about what are the key drivers of practice patterns and physician perspectives, recommendations for policy, etc. should all be removed, or possibly relegated to appropriately qualified comments in the Discussion.

We reiterate here our earlier responses regarding standards for qualitative methodology.However in response to this comment we have strengthened our justification for the connection between our recommendations and the empirical data; please refer to the highlighted text on p18-19. We have also edited the title of the manuscript to make this distinction clearer for all readers. Doctors' perspectives on PSA testing illuminate established differences in prostate cancer screening rates between Australia and the United Kingdom: A qualitative study.

Comparisons of incidence and mortality rates between the two settings leave out any differences in primary treatment and follow-up care, but these must be mentioned even if only to observe that they are similar (if this is the case).
Caution must indeed be taken when interpreting comparisons in incidence and mortality data. We have added the following text to the introduction (L102):
There are many reasons for variation in incidence and mortality rates, which could be due to

underlying differences in prostate cancer risk and population age structures, men presenting for testing, access and availability of treatment options, cancer coding and registration, and diagnostic processes (such as availability of PSA testing and improved diagnosis).

"Grounded theory", which may be unfamiliar to many readers, is mentioned at the start of the methods with no explanation. If this is retained at all, I suggest moving it to the Discussion where it should be clearly explained and the speculative nature of any commentary related to the small selective samples in this study should be explicitly stated.

We have added the following paragraph to the Methods section outlining the core principles of a grounded theory study. Note that grounded theory is an extremely well-established empirical methodology.
Grounded theory is a method of conducting qualitative research that focuses on creating conceptual frameworks or theories through building inductive analysis from the data (11). Grounded theorists are led by the experiences of the people in their inquiry and the substantive theories they develop closely reflect what those people experience and do. Specific methods of data collection and analysis are used to identify patterns in the research data. The twin foundations of grounded theory are the processes of constant comparison (a simultaneous and concurrent process of coding and analysis) and theoretical sampling (sampling with the aim of developing the properties of a developing idea or theory). These methods together guide the systematic development of emerging theory, and ensure findings remain firmly grounded in the collected data. All study authors have been formally trained in the methods described; SC has particular expertise in grounded theory methodology.

As stated above in response to an earlier comment, we have strengthened our justification for the connection between our recommendations and the empirical data: please refer to the highlighted text on p18-19.

What exactly is meant by symptomatic vs asymptomatic men? For example, a man with a first-degree family relative with prostate cancer may be more likely to request a PSA test and to receive it in both settings, yet in both settings he would be considered "asymptomatic", correct?

We agree with the reviewer that terminology regarding PSA testing is often difficult to interpret. This is in part because similar medical technology is used for both a screening test and a diagnostic test. This analysis was concerned with the use of the PSA test as a tool for screening only.
Screening tests are used to indicate or detect potential disease indicators (in this case, a risk marker – raised PSA) in people not presenting signs or symptoms for that disease (i.e. they are asymptomatic). ‘Diagnosis’ on the other hand is using a test to establish the presence or absence of disease as a basis for treatment decisions in symptomatic or asymptomatic individuals with a positive screening test (as a confirmatory test - in this case, a biopsy).
In response to this comment, we have included a brief definition of what we mean by ‘asymptomatic’ in the manuscript; L88.

In this paper, ‘asymptomatic’ will refer to those men attending clinical practice with no prior indications associated with prostatic disease. This is in contrast to the detection of prostate cancer in symptomatic men: men who have symptoms that could be related to locally advanced or metastatic prostate cancer such as frequency of urination, new onset bone pain and/or neurological symptoms involving the lower extremities (12).

We agree this is a grey area. A man with a strong family history of prostate cancer is still considered to be asymptomatic (i.e. as currently having no signs or symptoms of the disease) but high-risk. A GP might be more likely to consider PSA screening of these men, and perhaps at a younger age, as recommended by some medical bodies (as shown in our analysis).

However, having a risk factor does not mean a man will develop prostate cancer; conversely the absence of any risk factors does not protect a man from developing prostate cancer.
It is also very common for men over 50 years of age to present to primary care experiencing symptoms related to urinary flow, urgency, or control. These can be symptoms of early prostate cancer, but in most cases, these symptoms are caused by non-cancerous, enlargement of the prostate. Localised/early prostate cancer does not usually present as symptomatic; late stage/metastatic cancer on the other hand will often present as bone pain.
While we asked GPs to focus on care for men without symptoms, we also had open discussions with them about the complexities of prostate symptomatology and what they interpreted 'asymptomatic' to mean, which are reflected in the analysis.

1. Pope C, Mays N. Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. BMJ: British Medical Journal. 1995;311(6996):42.
2. Britten N. Qualitative research: qualitative interviews in medical research. Bmj. 1995;311(6999):251-3.
3. Mays N, Pope C. Assessing quality in qualitative research. British medical journal. 2000;320(7226):50.
4. Mays N, Pope C. Rigour and qualitative research. BMJ: British Medical Journal. 1995;311(6997):109.
5. Creswell JW. Qualitative inquiry and research design: Choosing among five approaches: Sage; 2013.
6. Morse J. Designing funded qualitative research. In Norman K. Denzin & Yvonna S. Lincoln (Eds.) Handbook of qualitative research (2nd ed. pp220-35) Thousand Oaks, CA: Sage; 1994.
7. Charmaz K. Constructing grounded theory : a practical guide through qualitative analysis. London: SAGE Publications; 2006.
8. Green J, Thorogood N. Qualitative methods for health research: Sage; 2013.
9. Williams N, Hughes LJ, Turner EL, Donovan JL, Hamdy FC, Neal DE, et al. Prostate-specific antigen testing rates remain low in UK general practice: a cross-sectional study in six English cities. BJU international. 2011;108(9):1402-8.
10. Moss S, Melia J, Sutton J, Mathews C, Kirby M. Prostate-specific antigen testing rates and referral patterns from general practice data in England. International journal of clinical practice. 2016;70(4):312-8.
11. Bryant A, Charmaz K. The Sage handbook of grounded theory: Sage; 2007.
12. American Urological Association. Early Detection of Prostate Cancer: AUA Guideline. 2013.

VERSION 2 – REVIEW

REVIEWER	Michael Kirby The Prostate Centre London
REVIEW RETURNED	28-Aug-2016

GENERAL COMMENTS	My comments adequately addressed

REVIEWER	Roman Gulati Fred Hutchinson Cancer Research Center, USA
REVIEW RETURNED	08-Sep-2016

GENERAL COMMENTS	The revision provides helpful clarifications about the qualitative study design. I believe the paper is substantially improved.

	However, I still have concerns that patterns observed in the two small selective samples in Australia and the UK are interpreted and presented as representative of general practice patterns. The study design does not support inference to general practice patterns. The authors explain that the grounded theory methods used in this study "ensure findings remain firmly grounded in the collected data," and I agree that descriptive summaries of their interviews is illuminating. However, numerous statements go beyond the data collected. For example, "Patients within the Australian marketplace model of healthcare are regarded as consumers," "In Australia, the blend of pro screening culture with mainstream media creating demand from patients for PSA tests directly creates a market for screening," "It seems likely from our data that GPs from Australia and the UK are following different 'mindlines', shaped by their respective cultures, contexts, and experiential knowledge," etc. Similarly, the study design does not support conclusions about what are the primary drivers of screening behavior or how policy should change to reduce screening. Again, descriptive summaries of interviews is interesting and hypothesis-generating, but they do not form a credible basis for policy recommendations. Finally, the authors state: "the diversity of views evident in our data suggests that it is very unlikely that our sample was biased towards a particular explanation or view of PSA testing." But this is a false assurance. A diversity of views does not imply an unbiased sample or an unbiased majority opinion, which forms the primary basis for the authors' comparisons. How many, how frequently, or how consistently an answer was given by GPs in these two small selective samples is useful to summarize these data but should not be assumed to be representative of national views.

VERSION 2 – AUTHOR RESPONSE

Reviewer 3.

I still have concerns that patterns observed in the two small selective samples in Australia and the UK are interpreted and presented as representative of general practice patterns. The study design does not support inference to general practice patterns. The authors explain that the grounded theory methods used in this study "ensure findings remain firmly grounded in the collected data," and I agree that descriptive summaries of their interviews is illuminating. However, numerous statements go beyond the data collected. For example, "Patients within the Australian marketplace model of healthcare are regarded as consumers," "In Australia, the blend of pro screening culture with mainstream media creating demand from patients for PSA tests directly creates a market for screening," "It seems likely from our data that GPs from Australia and the UK are following different 'mindlines', shaped by their respective cultures, contexts, and experiential knowledge," etc.

We have made the following changes to the particular points of contention raised by the reviewer.

• The GPs who participated in this study often spoke of patients as consumers, who maintained substantial individual choice in healthcare decisions.
• Australians have been shown empirically to have attitudes broadly in favour of cancer screening. The Australian media has been shown empirically to deliver a generally pro-PSA-screening message. These two combined seem likely to increase rather than limit patient demand for PSA testing, and thus to promote rather than retract a market for screening.

• We hypothesise from our data that GPs from Australia and the UK are following different ‘mindlines’.

Similarly, the study design does not support conclusions about what are the primary drivers of screening behavior or how policy should change to reduce screening. Again, descriptive summaries of interviews is interesting and hypothesis-generating, but they do not form a credible basis for policy recommendations.

We have changed the language from suggested additions to Australia’s push for reform to suggested areas for consideration and evaluation, rather than representing actual policy recommendations as the previous wording implied.

• In the following section we suggest areas for consideration and evaluation (alongside the NHMRC guidelines), which may potentially decrease use of the PSA test for screening purposes in Australian primary care.

Finally, the authors state: "the diversity of views evident in our data suggests that it is very unlikely that our sample was biased towards a particular explanation or view of PSA testing." But this is a false assurance. A diversity of views does not imply an unbiased sample or an unbiased majority opinion, which forms the primary basis for the authors' comparisons. How many, how frequently, or how consistently an answer was given by GPs in these two small selective samples is useful to summarize these data but should not be assumed to be representative of national views.

We have removed this statement from the limitations section.